Navy Collaborative Integrated Information Technology Initiative

(NAVCIITI)

ONR Grant N00014-00-1-0549

Report No. 14

Quarterly Report

Year 2

Submitted to:

Mr. Paul Quinn
ONR 311
Office of Naval Research
Ballston Centre Tower One
800 North Quincy Street
Arlington VA 22217-5660

Submitted by:

Virginia Tech
Blacksburg VA 24061

1 November 2000 – 31 January 2001

20010320 081

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY)	2. REPORT DATE	3. DATES COVERED (From - To)
30-08-2000	Technical	Nov 1, 2000 - Jan 31,2001

4. TITLE AND SUBTITLE

Navy Collaborative Integrated Information Technology Initiative

5a. CONTRACT NUMBER
N00014-00-1-0549

5b. GRANT NUMBER
N00014-00-1-0549

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Kenneth Reifsnider
Rick Habayeb

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Virginia Polytechnic Institute & State University
Office of Interdisciplinary Programs
306 Burruss (0244)
Blacksburg, VA 24061

8. PERFORMING ORGANIZATION REPORT NUMBER
14

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office ofNaval Research
Ballston Centre Tower One
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Quarterly progress report on NAVCIITI hardware and software configurations of smart antennas, multi-function antennas, secure configurable platform, command & control visualization, visualization and HCI, collaboration workspace, digital ships, mechanically flexible displays, fiberoptic LAN, network protocol interoperability, network system interoperability, and scheduling and resource management.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPONE NUMBER (Include area code)
U	U	U		100	540-231-9359

DISTRIBUTION STATEMENT AUTHORIZATION RECORD

Title: NAVY Collaborative Integrated Information Technology Initiative

Authorizing Official: Paul Quinn

Agency: ONR Ph. No. (703) 696-5753

☐ Internet Document: URL: _____
(DTIC-OCA Use Only)

Distribution Statement: (Authorized by the source above.)

☒ **A**: Approved for public release, distribution unlimited.

☐ **B**: U. S. Government agencies only. (Fill in reason and date applied). Other requests shall be referred to (Insert controlling office).

☐ **C**: U. S. Government agencies and their contractors. (Fill in reason and date applied). Other requests shall be referred to (Insert controlling office).

☐ **D**: DoD and DoD contractors only. (Fill in reason and date applied). Other requests shall be referred to (Insert controlling office).

☐ **E**: DoD components only. (Fill in reason and date applied). Other requests shall be referred to (Insert controlling office).

☐ **F**: Further dissemination only as directed by (Insert controlling DoD office and date), or higher authority.

☐ **X**: U. S. Government agencies and private individuals or enterprises eligible to obtain export-controlled technical data in accordance with DoD Directive 5230.25.

NOTES: _____

J Keith
DTIC Point of Contact

30 Mar 2001
Date

Task 1a.1.1: Wideband Vector Channel Measurements

This task focuses on the measurement of vector channels for wideband communication systems. These measurements provide an important (and previously absent) knowledge base for the modeling of wideband communication channels and the design of wideband wireless systems that employ smart antennas. In turn, this knowledge base will stimulate research in innovative diversity-combining algorithms and adaptive filtering techniques, including space-time adaptive processing technique.

The previous report on Virginia Tech's measurement receiver discussed the design, construction, and capabilities of the wideband vector channel measurement receiver. Since the development of this hardware, software has been written to implement radio receiver algorithms that operate on measured data. Software receiver modules have been developed that measure strength, delay, and phase of multipath components arriving at each element of a antenna arrays.

Using the FPGA transmitter described in the previous report, a 1023-chip pseudo random binary (PN) sequence modulates an RF carrier. This direct-sequence, spread spectrum signal is transmitted using a single antenna element. A recently constructed, four-element linear array is used at the receiver to receive the signals at the four-channel measurement receiver.

Figure 1 shows the hardware that downconverts the RF signals received at each of the four antenna elements to IF signals. Each IF signal is sampled at 1 Gsps and processed in software. The next figure illustrates a block diagram of the software that processes the IF signals. All filtering and baseband downconversion is performed by the software module. By executing a correlation between the received signal at each antenna port and a known PN sequence, a power delay profile is produced. The power delay profile represents the impulse response of the radio channels between the transmitter antenna and each of the receiver antenna elements.

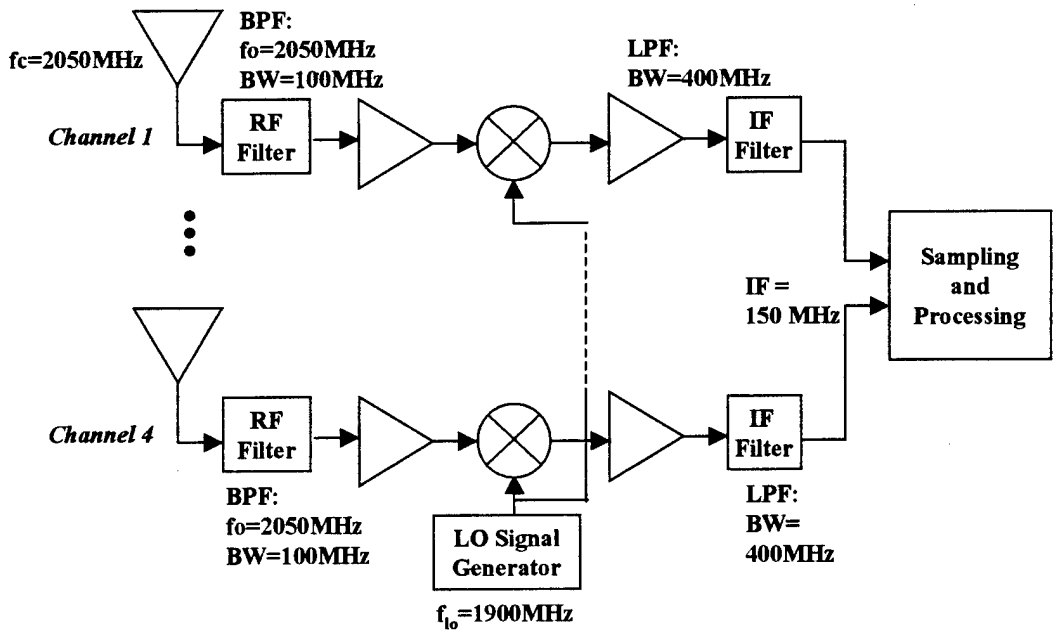


Figure 1. Block diagram of the receiver hardware.

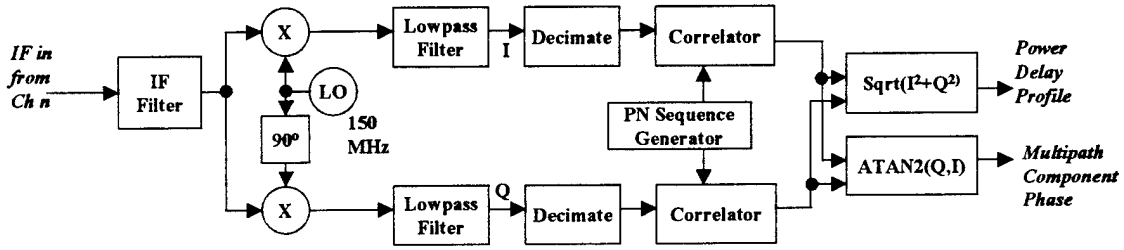


Figure 2. Block diagram of the software processing for the receiver

Illustrated in Figure 2 is a block diagram of the radio receiver functionality implemented in software. The software operates on the sampled IF signals from the receiver hardware. The output of the software is the power delay profile and multipath phase information. This block diagram is replicated for each receiver channel.

Figure 3 shows a sample power delay profile, measured in an indoor radio channel. The top of the figure shows the magnitude and delay of multipath components relative to the strongest-arriving component. This particular channel exhibited two strong multipath components within 15 dB of the line-of-sight component and arriving within 150 ns. Hundreds of power delay profiles like this one are collected and processed to compute statistical parameters of the radio channel that characterize time dispersion in the channel due to multipath.

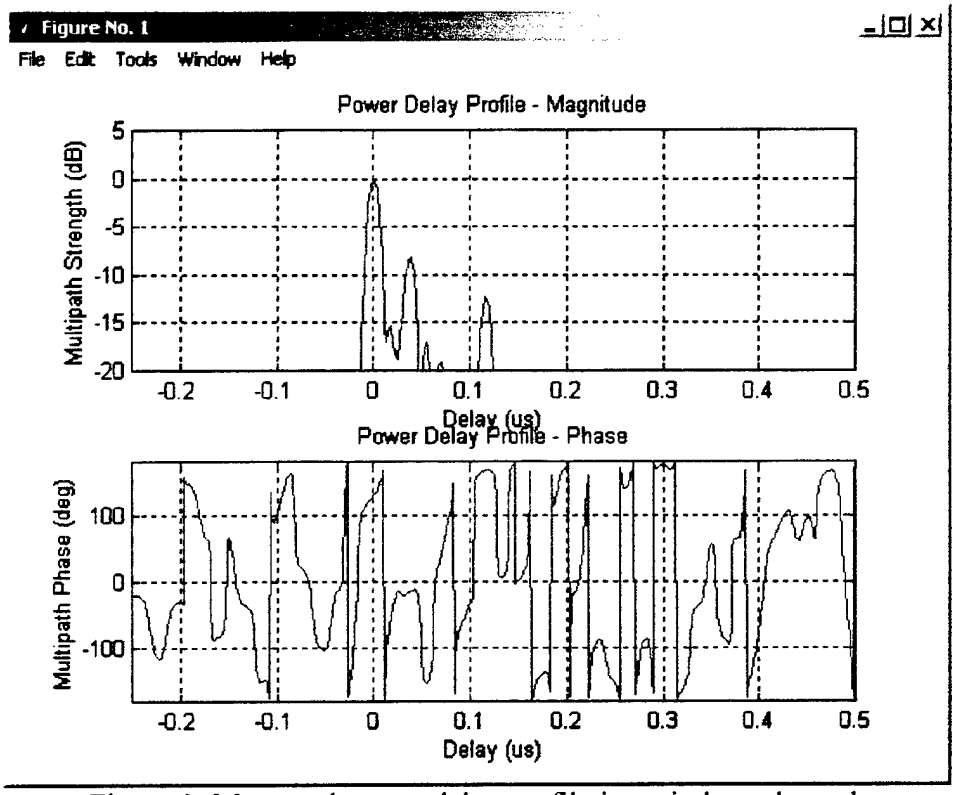


Figure 3. Measured power delay profile in an indoor channel.

The software and hardware developed for measuring multipath will be used to characterize and model radio channels in several environments. Preparations are being made to the measurement system to use the system in a mobile environment. The system will be operated in an automobile to measure and characterize real world radio channels (outdoor-outdoor and outdoor-indoor). Measuring and modeling radio channels will assist with simulating and testing antenna array algorithms (fully adaptive and diversity) and development of new adaptive algorithms, which may be able to utilize the statistical knowledge of the radio channel to improve performance of the antenna array algorithm.

Task 1a.1.2: Examine Transmit Diversity Strategies

This part of project will investigate transmit diversity techniques at the handset. Transmit diversity holds the potential of providing better link quality, improved battery life, and improved capacity. Our research shows that diversity implemented on the transmitter side provides better performance over a single antenna transmission system. Issues such as latency, complexity, and overhead of the feedback signal are fundamental limitations that are being explored. We are developing the algorithms for steering antennas on the handset with minimal feedback overhead and complexity. These algorithms will be tested using laboratory experiments and computer simulation.

During this quarter, effort was spent on working on developing a more sophisticated testbed and composing drafts for paper publications. The testbed development is towards implementing a multiple-input-multiple-output (MIMO) system as described in the Space-Time Coding section of this report. The system has multiple antenna elements at both transmitter and receiver and my focus is on the multichannel interface design for input/output (I/O) operations of this testbed. The interface design is intended for connecting multiple RF chains to the host DSP at the receiver, and connecting multiple signals from the host DSP to multiple RF transmitters.

Publications of the results from the research are also of prime interest for this quarter. Several topics have been identified from the preliminary report [1] and drafts have been composed in this regard. The topics include

1. Transmit diversity at the handset in a flat fading channel,
2. Real-time demonstration of wideband transmit diversity,
3. Transmit diversity measurements for an indoor channel,
4. Wideband channel measurements for an indoor and an outdoor-to-indoor channel.

The first paper presents transmit diversity techniques for a flat fading environment. This paper address proposed algorithms from convergence and simplicity of implementation point of view. The second paper describes a real-time hardware demonstration of a transmit diversity technique and discuss the testbed, transmit diversity technique and the measurement results. The third paper discusses narrowband transmit diversity measurements in an indoor channel. The fourth paper presents wideband channel measurements employing frequency sweep technique.

The activities include a seminar to the department of Electrical and Electronics Engineering at Bangladesh University of Engineering and Technology (BUET) during an overseas visit to Bangladesh. That seminar prompted interest among the faculty and students on various aspects transmit diversity.

Plans for the next quarter include completing the MIMO testbed and carrying out experiments with it. This will address applicability of transmit diversity for a wideband signal in different propagation environments.

There is an ongoing effort to build a wideband transmitter module and a multichannel receiver to conduct wideband channel measurements. The wideband transmitters are under construction. The receiver is based on a Gage Oscilloscope from Tektronix. The scope is a Deep Memory Oscilloscope or DMO that can capture up to 1GB of data. It has four channels and can be run at a sampling rate of 12-15 MHz using the external clock option. It will provide the capability to measure 3G like signal bandwidths. One effort will be to determine the feasibility of operating multiple transmitters that are synchronized both in time and phase, and implement transmit diversity with the transmitter array. The block diagram representation for this is shown in Figure 4.

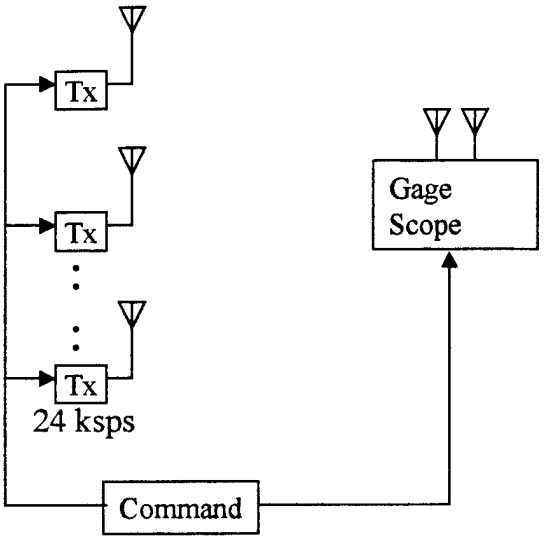


Figure 4. Block diagram to implement transmit diversity with the Gage scope.

Task 1.a.1.3: Space-Time Coding

A. Space-Time Processing

Research focused on comparing the performances of 2-D Rake receivers with the conventional Rake receiver and the conventional beamformer. The study included BER and capacity performances of a 2-D Rake receiver with varying number of antenna elements and Rake fingers. The performance tradeoff between antenna arrays and Rake fingers in various channel conditions was also studied. The performances of different beamforming techniques used in the 2-D RAKE receiver were also compared. A mathematical formulation for the output SINR of a 2-D Rake receiver was developed. Figure 5 shows the generic structure of a 2-D Rake receiver.

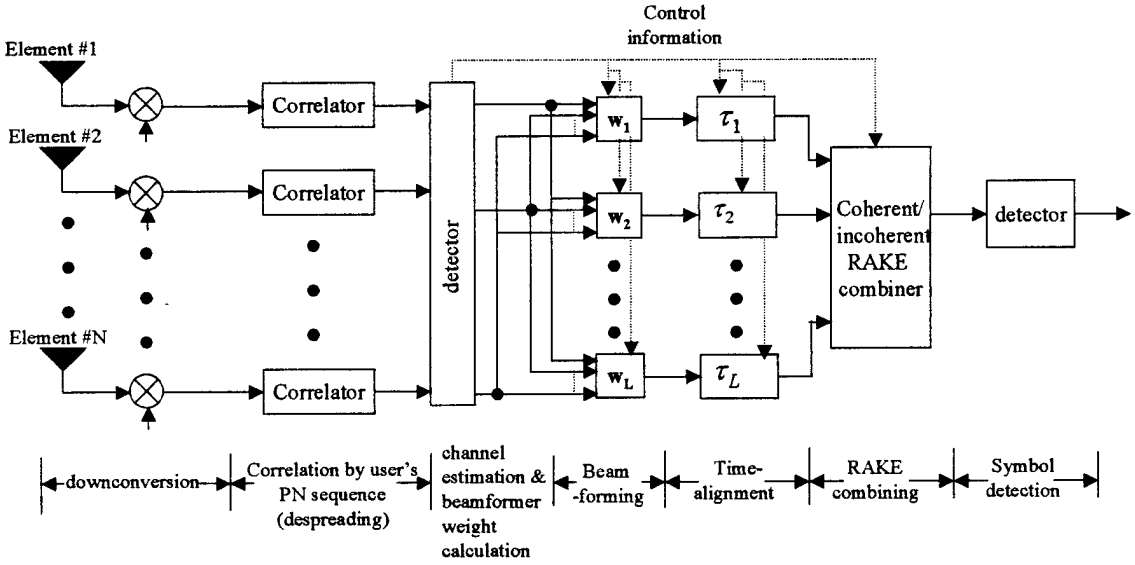


Figure 5. Block diagram of a generic 2-D Rake receiver.

The performance trade-off between the number of antenna elements and Rake fingers is illustrated in Figure 6. This is an interesting study that explicitly shows the relative importance of the number of elements in antenna array and the number of fingers in the Rake for a circular channel model. The receiver employed Recursive Least Squares (RLS) algorithm. The plots show that the 2-D Rake (4 element and 6 finger) performs better than 1-D Rake (1 element and 6 finger) and conventional beamformer (4 element and no Rake). It also shows that the 1-D Rake performs better than the conventional beamformer. It points out to the fact that more users can be supported if a system based on 1-D Rake or conventional beamformer is upgraded to a fully functional space-time receiver or a 2-D Rake receiver.

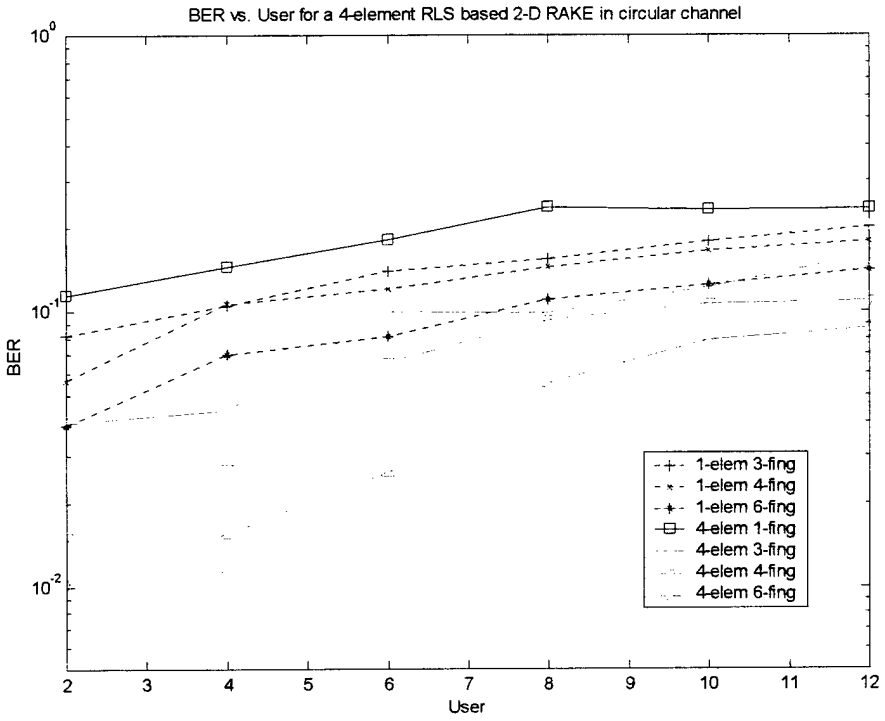


Figure 6. Performance comparison of a 2-D Rake receiver with 1-D Rake and conventional beamformer.

B.0 Space-Time Coding Test Bed

With the integration of Internet and multimedia applications in next generation wireless communications, the demand for reliable high data rate services is growing rapidly. Traditional wireless communications systems use a Single Input Single Output (SISO) channel, meaning one antenna on each side of the link. In a hostile environment containing high multipath, such as an indoor environment, these systems tend to be limited in data rate, or throughput, as well as in reliability. One strategy to mitigate these impairments is to introduce error-correction codes, which provide a form of time diversity. Space-Time coding combines the design of error-correction codes with

multiple-element antenna array technology, thus exploiting the spatial-temporal diversity inherent to the Multiple Input Multiple Output (MIMO) channel.

In this project, we design the Virginia Tech Space-Time Advanced Radio (VT-STAR) system that demonstrates the potential of the MIMO channel as compared to the SISO channel. The VT-STAR system is coupled with an MPEG-2 video stream to show a representation of the effect of the wireless channel on a video transmission in real-time. The block diagram of the hardware testbed is shown in Figure 7.

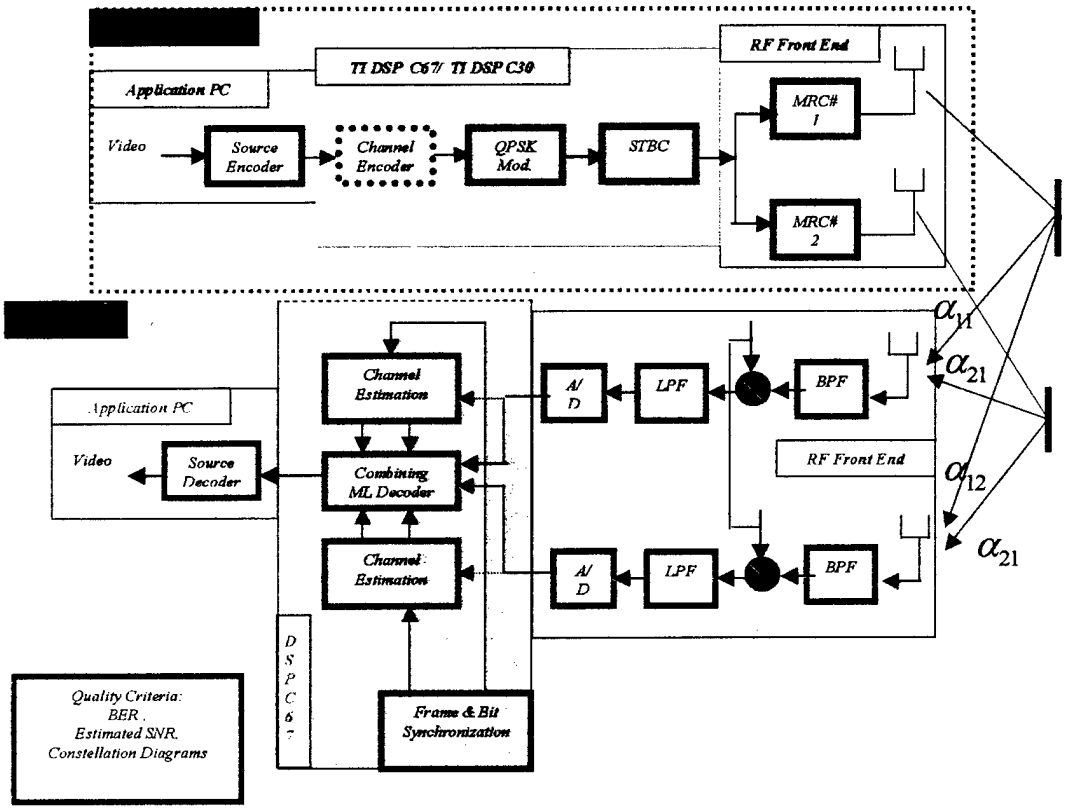


Figure 7. Block diagram for the testbed

The core algorithms are implemented on Texas Instruments (TI™) C6701 Evaluation Module (EVM) Bundles using TI™ TMS320C67 floating-point DSP processors.

Data conversion between the digital and analog domains is performed by TI™ THS56x1 EVM and TI™ THS1206 EVM for the transmitter and receiver, respectively. The radio frequency (RF) subsystem is composed of multi-channel transmitter and receiver chains implemented on hardware from multiple vendors.

The capabilities of the MIMO channel will be demonstrated in a Non Line of Sight (NLOS) indoor environment. Real-time monitoring of physical layer parameters, such as bit error rate and diversity advantage curves, as well as a video display will be presented on an attached personal computer.

This project presents a visual demonstration of the power of Space-Time Coding techniques. These techniques represent the leading edge of wireless research today, promising to deliver enhanced performance to the wireless communication systems of tomorrow.

B.1 Activities during the quarter

1. Design of a simulator code for Differential Space-Time Block Coding (D-STBC).
2. Comparison of differential vs. coherent STBC demodulation.
   - Frequency and Phase Errors.
   - Channel Estimation.
3. Derivation of analytic performance results for Space-Time Trellis Codes (STTC) operating in Rayleigh, Rician and correlated fading channels.
4. Design of a simulator code for symbol synchronization of a D-STBC scheme.
5. System level and algorithm design for space-time coding test bed.

B.2 Analytic Results

Generating function techniques have been extensively used to analyze the performance of both convolutional and trellis codes over AWGN and Rayleigh fading channels. We apply the generating function technique with the appropriate modifications in order to upper bound the bit-error probability of Space-Time Trellis Codes (STTC) over Rayleigh and Rician fading channels. We find the STTC, which combines the design of channel code with symbol mapping onto multiple transmit antennas, to be very similar to Multiple Trellis-Coded Modulation (MTCM). Following the theory developed for MTCM, we build the generating function of the STTC and combine it with the union bound to evaluate an upper bound on the bit-error probability. Next, we generalize this bounding technique to support the case of STTC with multiple receive antenna elements. As a study case, a class of STTC with 2 transmit antenna elements and multiple receive antenna elements is considered. For these codes, the upper bound is found to be inversely proportional to the Lth power of E_b/N_0, where L corresponds to the diversity advantage of the scheme. This diversity advantage is proportional to the number of antenna elements in both ends of the link and can be observed as the asymptotic slope of the error rate curves. Furthermore, comparing simulation results with the derived upper bounds as shown in Figure 8, we conclude that these bounds are as tight as 2-3 dB to the actual performance of STTC over Rayleigh fading channels.

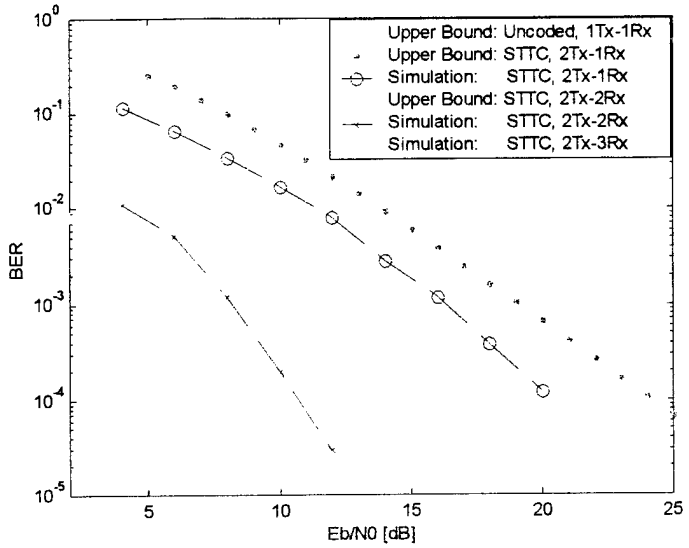


Figure 8. Error Bounds for Space-Time Trellis Codes over Flat Rayleigh Fading Channel

B.3 Publications

R. Gozali and B.D. Woerner, "Upper Bounds on the Bit-Error Probability of Space-Time Trellis Codes Using Generating Function Techniques," To appear in proceedings of VTC' Spring 2001.

R. Gozali and B.D. Woerner, "Performance Analysis of Space-Time Trellis Codes in Correlated Rayleigh Fading Channels," To be submitted to GlobeCom 2001.

R. Gozali and B.D. Woerner, "Performance Analysis of Space-Time Trellis Codes in Rayleigh, Rician and Correlated Fading Channels", To be submitted to *IEEE Transactions on Communications*.

B.4 Plans for Next Quarter

By next quarter we will have the DSP implementation of some of the initial space-time coding algorithms. We will have an initial prototype and will have begun testing of the system.

Task 1a.1.4: Channel Modeling and Simulation

Hidden Markov Model (HMM) that has proved itself to be very successful in the field of speech recognition, artificial intelligence, machine learning, and image and handwriting recognition, can also be used to model error sources in communication channels and more specifically to model wireless channels.

Modeling Hidden Markov Model is the method to estimates its parameters and we estimate the parameters of Hidden Markov Model for fading channels. We assume channel to be having many states and these states are analogous to the quantization

levels. Many algorithms are available to estimate the parameters of HMM and among them the Baum-Welch algorithm proved to be the most accurate and efficient.

The Baum-Welch algorithm to estimate the parameters of the HMM has been developed and now we are in the process of validating the model. The validation of the Hidden Markov Models is an important issue as it tells us that what are the limitations in using these models. We will compare the probability density function (pdf) generated from the State Transition Matrix of the HMM and the original pdf generated from the Wideband Code division Multiple Access (WCDMA) simulator.

During next quarter we will focus on answering the following questions.

- *How many states must be included in the HMM in order for results to be useful within the IMT-2000 standard?*
- *How robust are the models? In other words, under what conditions can the small changes in the waveform model give large changes in the HMM?*
- *How does state transition matrix converge as a function of simulation runlength? How accurate are the resulting models for simulations of a given runlength?*

In addition, we will try to determine model parameters based on the channel measurement (Viper). There are two ways to do that

- *Use Viper to develop power-delay profiles corresponding to a given test scenario. The resulting power-delay profiles can than be used in a waveform level simulation of the WCDMA system with the resulting error statistics used to parameterize the resulting HMM*
- *Use Viper to actually measure the error statistics on the channel.*

Reference:

[1]. Raqibul Mostafa, "Feasibility of Transmit Diversity at the Handset," a Preliminary Report submitted to the Ph. D. Committee, December 2000.

RESEARCH PROGRESS FOR THE THIRD QUARTER OF YEAR 2 FOR NAVCIITI PROGRAM ELEMENT 1.2,

Wideband Antenna Research

Warren Stutzman, William Davis, Randall Nealy, Carey Buxton, Ko Takamizawa, and Seong-Youp Suh

Antenna Group
Electrical and Computer Engineering Department
http://antenna.ece.vt.edu

This report describes progress during the third quarter (1 Nov 2000 – 31 Jan 2001) of Year 2 on the multifunction antenna research portion of the NAVCIITI program. Our work is in three areas: development of measurement and analysis capabilities for wideband antennas, research on wideband antenna elements, and research on wideband antenna arrays. This report discusses each of these areas. We have continued our industrial involvement with Harris Corp. and Northrop Grumman in the design, analysis, and testing of wideband array antennas. The activities of both of these companies involve wideband array programs that include Navy activities.

Several aspects of the operation of large arrays have been identified through this research and not previously highlighted in the literature. These efforts have identified further work with the potential of substantial improvements for wideband array systems, typical of many systems used by the Navy.

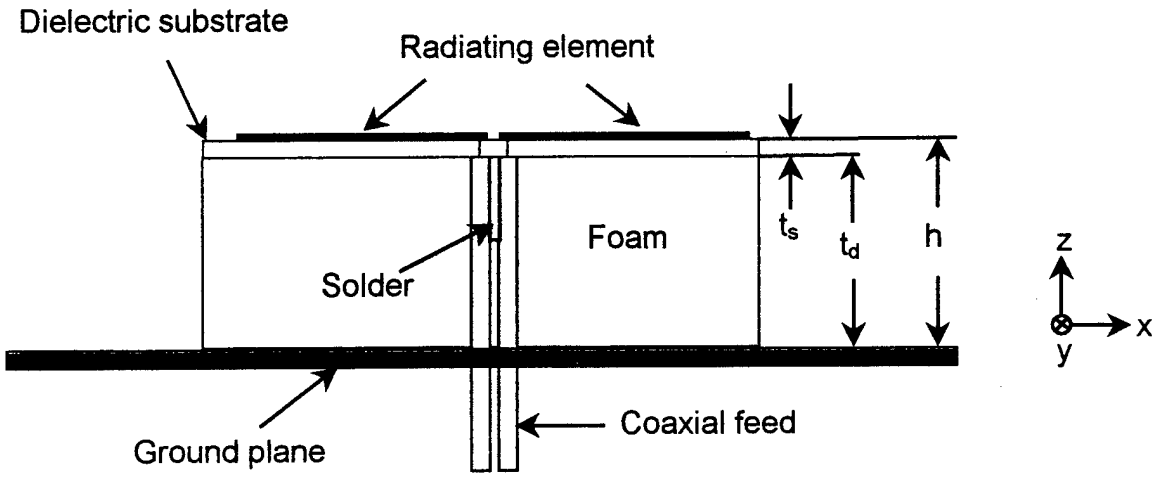
Task 1.2.1 Completion of the Test Facility

The spherical near field scanner and network analyzer equipment was installed and has been used for a variety of measurements, including measurements on the antennas to be discussed in the following tasks. Major efforts in this task has been to verify the performance of the near-field range using known antennas as well as standard gain horns. Future plans for work on this task include field plots for the interior of the chamber and characterization of the performance and limits of operation.

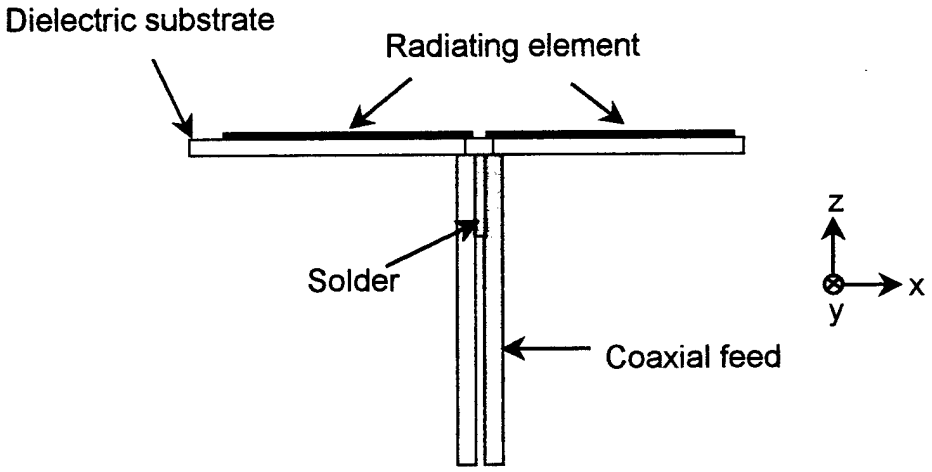
Task 1.2.2 Wideband Element Antenna Design

*** Following section contains patentable material. It shall be kept confidential.**

The emphasis of the current quarter was the investigation of the effects of the ground plane on Fourpoint antenna elements. The Fourpoint antenna was designed for operation between 6-10 GHz. The side view of the antenna is shown in Fig. 1.2.2-1. The impedance and radiation patterns were computed for two cases, with and without a ground plane using a commercial electromagnetic FDTD code, Fidelity. The ground plane was positioned below the radiating element with a $\lambda_c/4$ separation, where λ_c is the wavelength at the center frequency.



(a) With ground plane

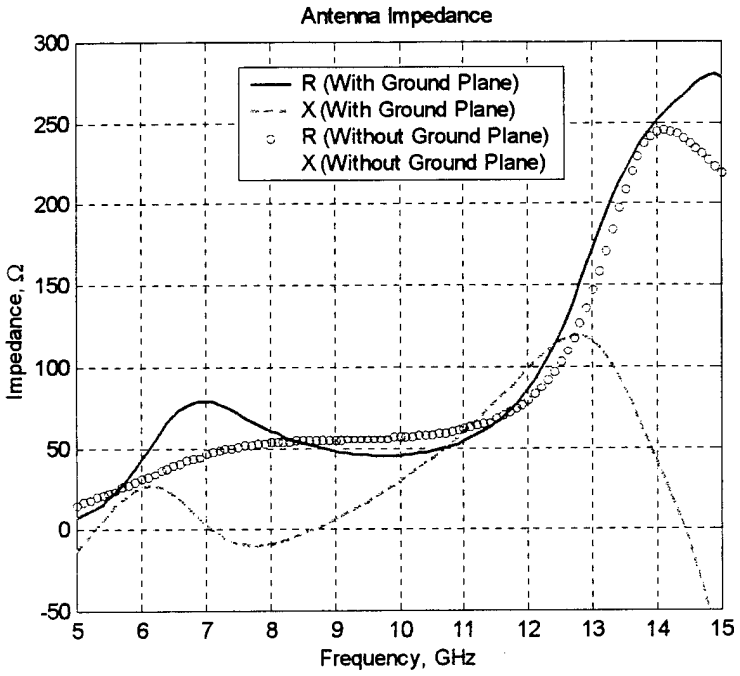


(b) With ground plane

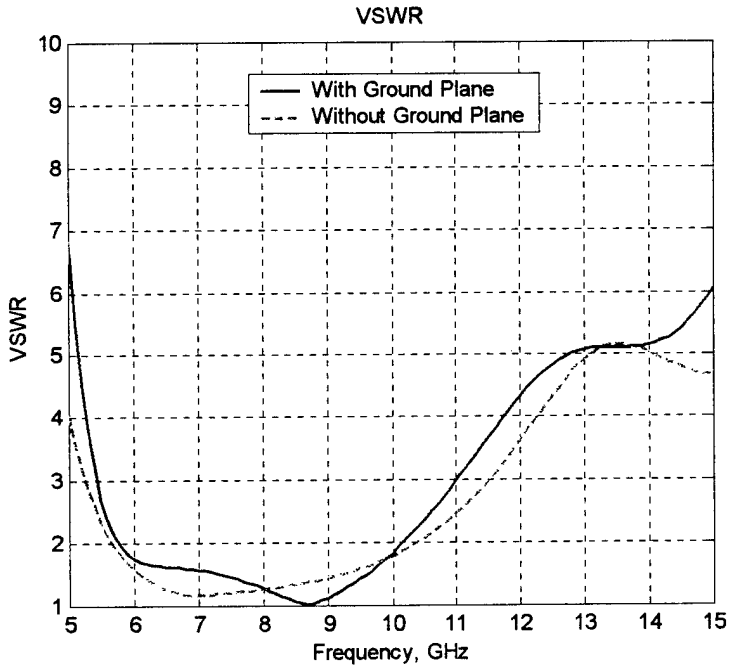
Figure 1.2.2-1 Sideview of the Fourpoint antenna

Fig. 1.2.2-2 shows the computed impedance and VSWR curves. The impedances are close for both cases over the band except at low frequencies. The impedance of the Fourpoint antenna without ground plane increases monotonically at low frequencies rather than the abrupt increase for the Fourpoint antenna with a ground plane. Though the computed results are not presented graphically, it was found that a small separation between the radiating element and the ground plane causes a more abrupt overshoot and undershoot in the impedance at low frequencies. In other words, the ground plane has a tuning effect on the antenna impedance at low frequencies.

The ground plane also directs the radiation pattern to broadside (z direction) by reflecting the radiated energy upward so that it increases the antenna gain. Fig. 1.2.2-3 shows the computed gain patterns in the E and H planes. The patterns were compared for two cases, with and without ground planes. About a 6 dB antenna gain improvement was achieved with the ground plane.

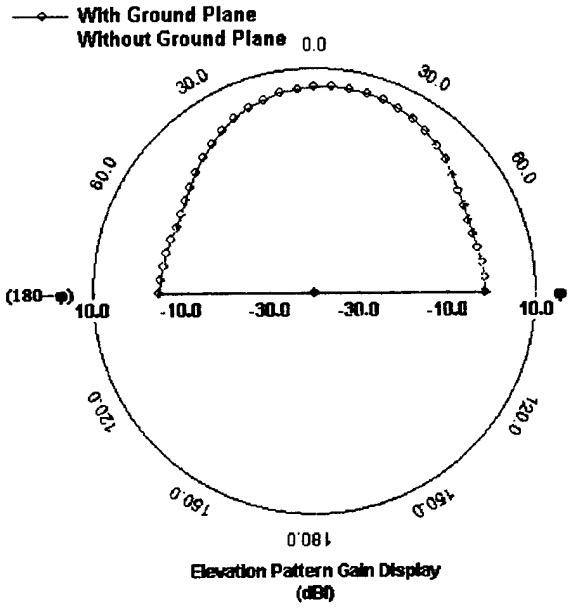


(a) Antenna impedance comparison

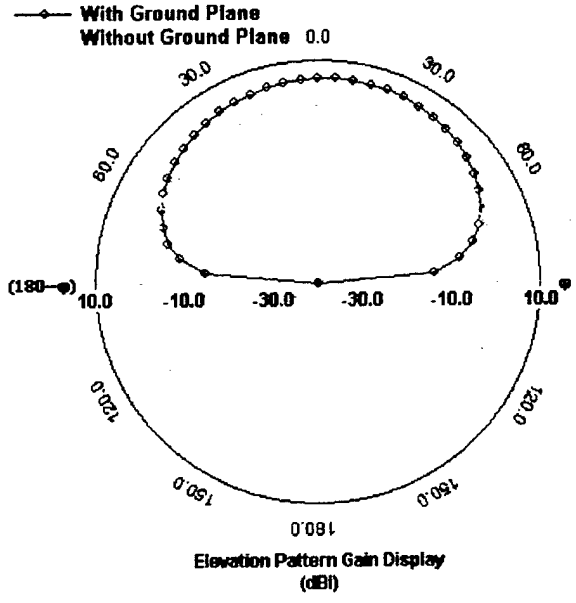


(b) VSWR for 50 Ω

Figure 1.2.2-2 Computed antenna impedance and VSWR comparison for with and without ground plane in the Fourpoint antenna.



(a) E-plane gain patterns



(b) H-plane gain patterns

Figure 1.2.2-3 Computed gain patterns for with and without ground plane in the Fourpoint antenna.

Task 1.2.3 Wideband Element Antenna Measurements

A Fourpoint antenna (with ground plane) operating at 6-12 GHz was constructed and measured at Harris Corp. We are teamed with Harris on an extensive development program for both Foursquare and Fourpoint elements and arrays. A hybrid coupler was connected to the input to measure the VSWR shown in Fig. 1.2.3-1. The hybrid data includes the measured S11 data at the actual hybrid input and the ideal hybrid data derived from measured S-parameters at both input port without the hybrid. Note the excellent agreement, verifying the measurement process. The VSWR curves are well matched and are under VSWR=2.5 from 5.5 to 16 GHz which is about a 98 % impedance bandwidth. This is the largest bandwidth achieved to date.

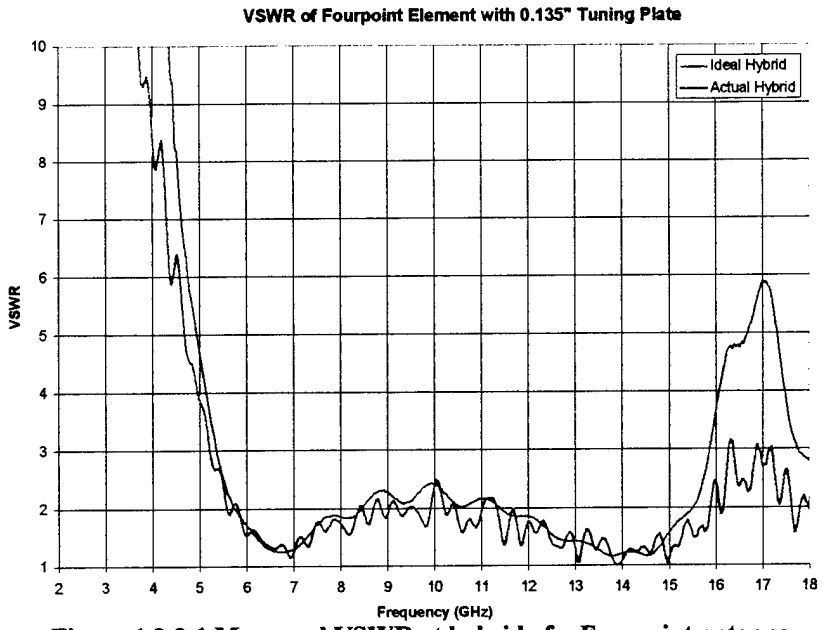


Figure 1.2.3-1 Measured VSWR at hybrid of a Fourpoint antenna.

Task 1.2.4 Extended Bandwidth Elements

As a result of an extended bandwidth element study, the Fourpoint antenna was invented. The Foursquare antenna is a predecessor of the Fourpoint antenna. An extensive parametric study is underway to lead to the development of more extended bandwidth elements.

Task 1.2.5 Array Antenna Investigations

The performance of large array antenna systems is typically considered to be limited by the effect of mutual coupling between the elements that vary with frequency and scan direction. The coupling effects are often stronger for wideband array antennas. The performance characteristics of a large array are usually predicted using a combination of infinite array antenna analysis techniques such as waveguide simulation [1] and Floquet series boundary cell analysis [2], along with active element pattern analysis [3] and network analysis. A waveguide simulator can be implemented with either experimental or numerical simulations. However, waveguide simulator techniques are applicable only for certain geometries, frequencies and scan angles. A Floquet series analysis is more general than waveguide simulator technique. It can be applied to an infinite linear or an infinite planar array of any uniform geometry. Unfortunately, the Floquet expansion method can only be implemented as a numerical simulation. There is no equivalent experimental method to verify the numerical analysis.

Sometimes, an antenna element in an array environment is evaluated using the active-element input impedance. The active-element input impedance of an element is determined by terminating all but the single fed element with a reference impedance. The reference impedance generally should be the equivalent source impedance for the array. Typically, 50 ohms is used for the reference impedance. The active input impedance is not a good indicator of the antenna element performance in the fully active array since the induced currents from neighboring elements modify the input impedance in a fully active array.

The fully active input impedance is the input impedance of an element in an array when all elements are excited. It is a function of array geometry and array excitation. In the case of an infinite array, all elements are excited uniformly in amplitude with a phase progression among the elements to produce the desired beam scanning. It has been shown that approximately a 9x9 or larger array fully active input impedances behave similar to a corresponding infinite array [4]. Unfortunately, it is difficult to numerically simulate or experimentally evaluate 9x9 arrays.

We have developed a novel technique to estimate fully active input impedance of infinite array antennas. The technique is based on an eigenvalue analysis of the coupling matrix. In this technique, the antenna characteristics of infinite array are predicted from the coupling matrix of a small subsection of the infinite array using eigenvalue analysis techniques. The eigenvalue analysis can be applied to both simulated data and measured data.

As an example, consider a linear array of parallel strip dipoles suspended over an infinite ground plane as illustrated in Fig. 1.2.5-1. The infinite array fully active input impedance was computed using in-house code ASIA [6]. For the case of a nine-element array, the mutual coupling coefficients were computed using IE3D written by Zeland [5].

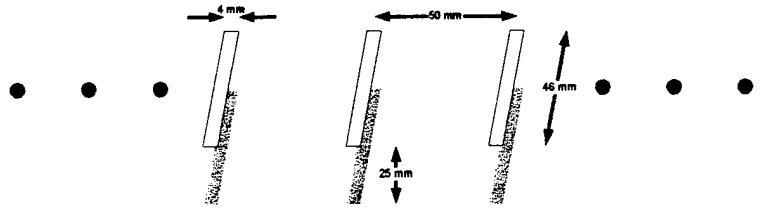


Figure 1.2.5-1 Illustration of an infinite linear array of parallel strip dipoles suspended over an infinite ground plane.

Fig. 1.2.5-2 shows a comparison of the fully active input impedance of an infinite array, the fully active input impedance predicted from the eigenvalue analysis of the nine element array, and the active element input impedance of the center element in the nine element array. The figure shows that the impedance predicted from the eigenvalue analysis compares well with the impedance of infinite array. It is also shown that active element input impedance is clearly different from the fully active impedance of array. In fact, the active element input impedance is actually closer to the input impedance of single isolated element than the fully active impedance.

Fig. 1.2.5-3 shows the comparison of fully active input impedance of infinite array to the impedance predicted from 2 elements, 3 elements, and 5 elements array using eigenvalue analysis. The figure shows that only 3 element coupling data is necessary to predict the impedance behavior of fully active infinite array for this particular geometry. This is a significant result since a design of a large array can be implemented using analysis of an array with three elements. Also, a considerable reduction in the complexity of experimental verification can be accomplished using this method.

In the future, the verification of the eigenvalue analysis technique will be applied to Foursquare array data. In addition, an extension of the eigenvalue analysis technique to scanned beam cases will be considered.

[1] A. A. Oliner and R. G. Malech, *Microwave Scanning* Antennas, R. C. Hansen, editor, Peninsula Publishing, Los Altos, Ca, *Volume III. Array Theory and Practice*, Chapter 3 "Mutual Coupling in Infinite Scanning Arrays," pp. 322-333, 1985

[2] R. J. Mailloux, *Phased Array Antenna Handbook*, Artech House, Boston MA, 1994.

[3] D. M. Pozar, "Active element pattern," *IEEE Trans. Antennas Propagat.*, vol. 48, pp 1176-1178, Aug. 1994.

[4] D. M. Pozar, "Analysis of finite phased arrays of printed dipoles," *IEEE Trans. Antennas Propagat.*, vol. AP-33, pp. 1045-1053.

[5] *IE3D Users Guide*, Zeland Inc.

[6] J.W. LaPean, W. L. Stutzman, W. A. Davis, "A computational tool for large planar phased arrays – ASIA," 1996 *Antennas and Propagation Society International Symposium*, AP-S Digest Vol 2, pp. 882-825.

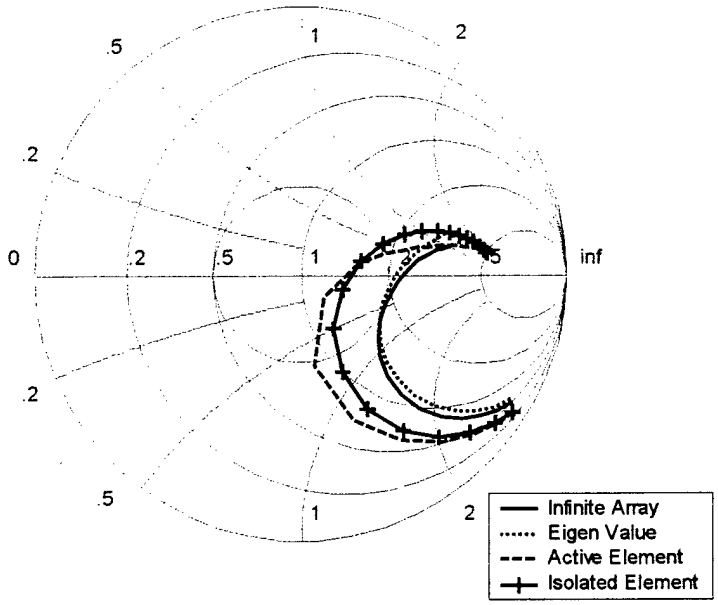


Figure 1.2.5-2. Comparison of fully active input impedance of infinite array to the impedance predicted using eigen value analysis, active element impedance, and isolated element impedance for the parallel strip dipole array shown in Figure 1.2.5-1. The eigen value analysis is applied on five element array.

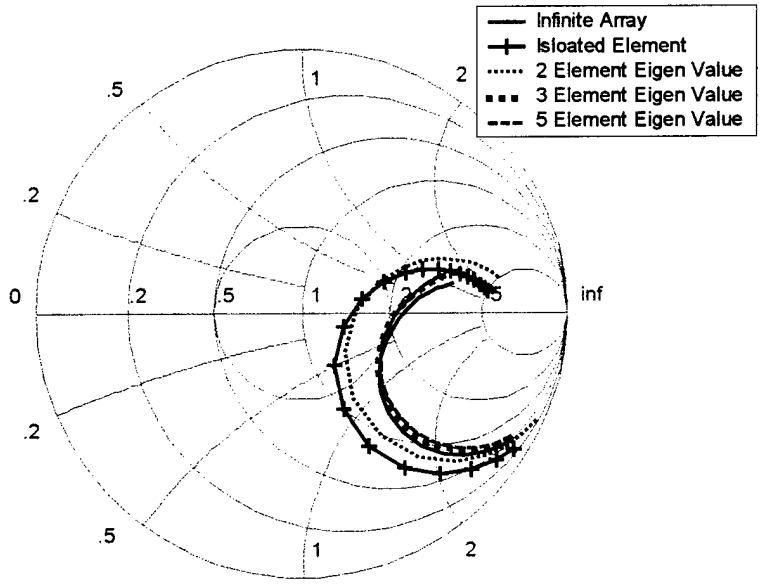


Figure 1.2.5-3 Comparison of fully active input impedance of infinite array to the impedance predicted using eigen value analysis applied to 2 element, 3 element, and 5 element array data.

Task 1.3:
Secure Communications Project Report for the 3rd Quarter of Year 2

Project Overview

The purpose of this project is to demonstrate the viability of a potentially highly adaptable low-cost technology for naval information management and communications systems. It is the intent of this project to demonstrate an alternative means of collecting, processing, storing, and transmitting information securely throughout the next-generation seaborne naval platforms. The technology featured in this task has been chosen based upon the following criteria:

(1) The focus of this project is on information management. The processing power of the underlying platform should be sufficient to address a variety of demanding computational tasks.
(2) The processing platform should utilize commodity (high production volume) components. The intent here is to reduce per-unit costs, manufacturing cost, and time-to-deployment as new devices become available.
(3) The hardware technology should be multifunctional, and be capable of fulfilling the operations of perhaps a variety of tasks with little or no modifications. The intent of this is to derive economy-in-scale by having one deliverable satisfy possibly several requirements.
(4) The system should handle data in a secure manner. System resources should be protected from unintentional (friendly misuse of system resources) or intentional (protection of software and hardware against theft and reverse engineering) threats.

Fourth Quarter Plans

In the course of the fourth quarter of 2000, we planned to develop a working prototype of a secure radio system. A basic application was also to be developed, demonstrating radio initialization and use.

Progress

A prototype radio system was constructed, consisting of a configurable base station that communicates with standard external transceivers. The system uses the Information Science Institute SLAAC configurable computing platform as a processing basis and Digital Wireless Corporation WIT2410 2.4GHz Spread Spectrum Wireless Industrial Transceivers for communication.

A talk application was embedded in the system to provide a demonstration of radio functionality. Using the European Space Agency LEON RISC core in the base station, the application allows users to communicate via a VT100 text-based interface. A username/password combination allows the base station to identify users on remote transceivers. The base then provides the remote device with user-specific data traffic.

Simple Exclusive-OR encryption was added to the configuration data upload path to provide a demonstration of configuration security. This operation defines an interface that simplifies the inclusion of other security measures in the system. This interface is currently being used to port the Blowfish encryption algorithm to the system.

At the end of the fourth quarter, the construction of a more functional prototype was discussed. The team decided to pursue an option that will demonstrate secure independent configurable receivers. Team goals were refocused to aggressively pursue this enhanced prototype using the technology developed in the initial prototype system.

Activities

Representatives of the Boeing Company Space and Communications group visited our facilities in mid-November. They were presented with a brief introduction to the Secure Communications Project.

In December, a presentation was made to the Virginia Tech Configurable Computing Group. A departing member of the Secure Communications Project team formally presented his work regarding RISC processor embedding to the group. The presentation included the first public demonstration of the working Secure Radio prototype.

Future Plans

In the course of the upcoming quarter, we plan to enhance the secure radio prototype. The updated prototype will provide secure handling of sensitive data in addition to secure hardware reconfiguration. This new system will be able to distinguish between three user classes and provide data appropriate to the class of the current user. More robust security of hardware configuration data streams will also be incorporated.

Issues

There are no outstanding issues.

Project 2.0 Visualization HCI and Collaboration
- Task 2.1: Command and Control Visualization
SOW 2.1.2: Design CAVE displays to interpret NUWC acoustic model results December 00
SOW 2.1.3: Demonstrate CAVE displays to interpret NUWC acoustic model results, April 01
SOW 2.1.4, Evaluate and modify CAVE display interfaces for NUWC acoustic model, July 01

Background: Our objective is to provide a distributed collaborative network of graphical and device independent tools in a shared virtual environment which can be used by Command and Control (C&C) personnel to gain a strategic advantage. Specifically we focus on the mission critical C&C interpretation of acoustic undersea data from towed arrays for the Naval Undersea Weapons Center (NUWC) using the CONRAY simulation models. These simulation models can be extended to "real-time" data acquisition systems. Under the direction of personnel from NUWC and the Naval Research Laboratory (NRL) we have identified a working prototype which we have successfully incorporated into our Device Independent Virtual Environment Reconfigurable-Scalable-Extensible (DIVERSE) tool that works in stereo in the (C)AVE Automated Virtual Environment (CAVE), Immersive Work Bench (IWB), Immersive Desk (I-Desk), desktop workstation simulator, and Head Mounted Display (HMD) systems at the Virginia Tech Center for Virtual Environments and Visualization (CVEV). Hence the idea of "DIVE" (Device Independent Virtual Environment). The DIVE in DIVERSE provides the basis for collaborative C&C.

Discoveries, Accomplishments, Test Results: Since the last quarterly report our Points of Contacts (POCs) at NUWC and NRL met at Virginia Tech November 13 and 14, 2000 for a demonstration of the DIVERSE-based CONRAY model and to begin discussion on the "next step" in the development of the CONRAY model. In correspondence with NRL, NUWC generated a task breakdown. CONRAY was rewritten and the code architecture redesigned to improve rendering speed and frame rate, to limit dependencies between CONRAY scene-building and the DIVERSE scene-rendering and navigation. A redefinition of subtasks within SOW 2.1.3 reflects a collaborative development effort between NUWC, NRL and Virginia Tech (VT). The new Command & Control (C&C) software visualization project has been labeled "MIX". The effort put in redesigning the DIVERSE version of CONRAY allowed for a very fast incorporation of the code into MIX.

Activities during quarter: On February 5, 2001 NUWC, NRL and VT agreed to coordinate efforts into the following task break down.

NUWC: (POCs Ken Lima, Ann Silva, Lauren Mathews, Richard Shell)
1. Develop a geometric method for determining points of intersection for complex conical angles.
2. Develop 3-D Eigenray Manifold 'ray trace' algorithms that employ 3-D reflection, sound-velocity profiles and propagation loss models.
3. Create multiple scenarios that will test operator effectiveness.
4. Make 3-D displays equivalent to Mark II block 1C combat system displays for 2-D vs. 3-D comparison testing:
   - Determine what combat system information should be displayed on the 3-D canvas
   - Update and finalize the form of the information bezel
5. Set up software to read in common navy databases.
6. Coordinate NUWC and NRL efforts and interact with VT as required.

NRL: (POCs: Larry Rosenblum, Robert King)
1. Integrate the information bezel with the DRAGON software - bezel should turn on and off from keyboard or I/O device.

2. Ensure control devices operate properly with all software and scenarios - adjust control device parameters as required.
3. Integration of NUWC code into the DRAGON software.
4. Documentation of DRAGON software
   - Diagram of DRAGON component interaction.

- Quick guide to code modification of elements pertinent to visualization
- Organized copy of source code
- Dragon presentation at NUWC

VT: (POCs: Ron Kriz, John Kelso, Fernando das Neves)

1. Explore optimization techniques to permit the code to run faster
2. Explore optional enhancements to the current code for incorporation into the DRAGON software such as:
   - Create a bathymetry contour following grid that maps to the bottom
   - Create bottom following vessel tracks.
   - Create the ability to lay generic texture maps of information such as navigational charts, bottom type, gravity maps, etc. on bathymetry.
3. Explore alternative devices and interaction techniques to improve selection and analysis of data subsets.

Ron Kriz and John Kelso visited Robert King at NRL on February 7. Rob showed us his progress on his part of the collaboration. Preliminary results suggest that NUWC, NRL, and VT now have a working relationship where outcome is based on technical issues that each organization has demonstrated expertise. DIVERSE Version 1.01 now includes bug fixes and a Dynamic Shared Object (DSO) for a Head Mounted Display (HMD). This satisfies a long term requirement to scale C&C software across heterogeneous Virtual Environments devices.

Plans for Next Quarter:

- *SOW 2.1.3: Demonstrate CAVE displays to interpret NUWC acoustic model results, April01*
- *SOW 2.1.4, Evaluate and modify CAVE display interfaces for NUWC acoustic model, July 01.*

NUWC will coordinate NUWC and NRL efforts and interact with VT as required. VT will maintain regular communications with NUWC regarding the development of these new subtasks. SOW 2.1.4 will require that the CAVE be operational before July. Presently the CAVE floor with motion platform is completed and construction of the CAVE will begin February 12, 2001. The CAVE will be operational late March. In the interim VT will develop on the I-Desk, IWB, and desktop workstations.

Outstanding Issues: Now is a good time to begin planning the integration of SOW 2.1.5*: Design Digital Ship CAVE interface (DCSI) for simulation of ship under fire, June 01.* From our experience this year with NUWC and NRL, software integration requires not only the existence of a Digital Ship Lab (DSL) Application Programming Interface (API), but DSL-API documentation as well; both are critical for planning this SOW.

Successful development of a working Command & Control (C&C) model, "MIX", between NUWC, NRL and VT has resulted in raising new questions in the development of C&C models that scale across heterogeneous VE systems. At the February 7, 2001 meeting at NRL we discussed issues related to future C&C models and concluded that Performer based code such as DRAGON and DIVERSE must be changed to OpenGL . But OpenGL presently lacks a scenegraph such as Inventor/VRML. We at VT believe these are also important issues that will need to be addressed in NAVCIITI year-3 C&C projects, if these same models are to prove useful to the current C&C community. From this collaboration it is clear that future C&C software architecture must be determined by hardware and software technical issues and limitations.

Task 2.2

Background

Virginia Tech has a long-term collaboration with NRL (Dr. E. Swan, Dr. J. Templeman, et al.) to research human-computer interaction (HCI) issues in virtual environments (VEs). Our recent work, funded by ONR and NRL and presented at the VR'99 Conference (the premiere international VR conference), was awarded "Best Technical Paper" of the conference by popular vote of conference attendees. Our work is conducting empirical studies to produce generalizable guidelines for designing VE user interfaces. The studies are using a map-based battlefield visualization VE called "Dragon." This prior and on-going work places us in an especially advantageous position to uniquely contribute to the proposed Digital Ship Lab (DSL) integrated testbed that is the "glue" for this NAVCIITI research.

We are using the Dragon software to explore use of eye tracking in VEs, producing and expanding guidelines for VE user interface design. To our knowledge, eye tracking has not been incorporated into a CAVE or an Immersive Workbench (IWB) elsewhere, much less systematically studied within these VEs. These issues are especially critical for (human-centric) design and evaluation of our Digital Ship concept. The objective of our current work is to concurrently explore two different facets of eye tracking:

1. use of eye tracking for data collection for usability analyses of VEs, and,
2. use of eye tracking as a multi-modal interaction technique in VEs.

Planned Effort for the Quarter

The planned effort for this reporting quarter (as stated in the most recent quarterly report – Report No. 13) included the following:

- Receive, install, and perform quality-assurance-check of upgraded ISCAN eye tracking systems.
- Install DIVERSE on the IWB and CAVE (equipment permitting).
- Install and refine the Dragon application on IWB and CAVE
- Finalize user interaction bugs with the Dragon application, and begin integrating multi-modal interaction components.
- Further develop evaluator's view using DIVERSE graphics. See http://www.diverse.vt.edu for more information on DIVERSE.

Other efforts performed this quarter (but not included in the most recent quarterly report) include:

- Further developing eye tracking DIVERSE services so that they may be included in the next official release of DIVERSE

We successfully accomplished all planned efforts during the quarter. Details of these activities are given below.

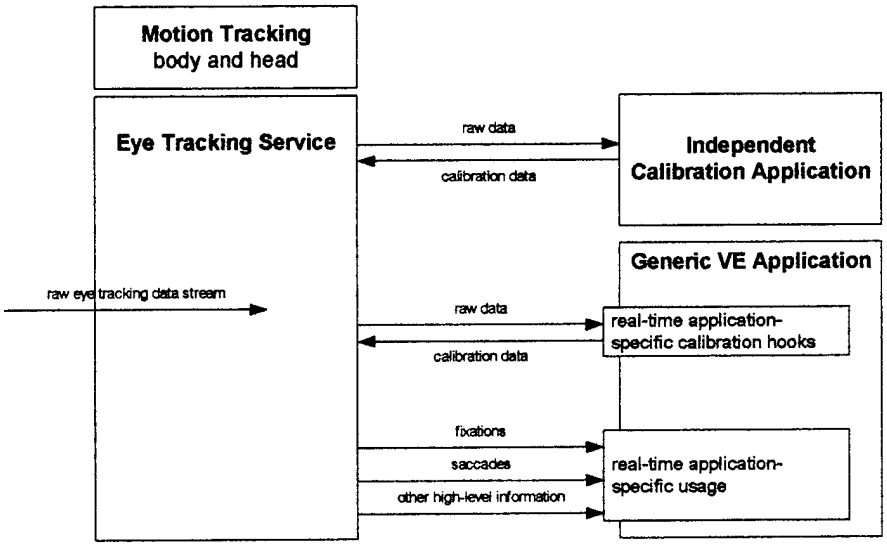
Accomplishments for the Quarter

Accomplishments for the quarter are presented below, and are grouped according to the original proposed NAVCIITI Year 2 tasks.

Integrated: the first "I" In NAVCIITI

One of the most significant accomplishments of the quarter includes our software development efforts that have produced a refined eye tracking service which is completely compatible and integrated into DIVERSE. That is, our efforts have produced an integrated product jointly developed by our NAVCIITI team and other researchers in NAVCIITI. Our eye tracking service is expected to be included in the next major release of DIVERSE.

The eye tracking DIVERSE service provides a communication protocol and connection agreement between an ISCAN eye tracking system and any DIVERSE-compatible computer system. It incorporates information from the head tracking system so that all eye-based information is coordinated with a user's body and head movements. A component diagram of how the eye tracking service coexists with VE-based applications (e.g., DRAGON) is presented below.



An example of how the eye tracking service is used by generic VE application.

We have increased the speed of communication between the eye tracking system and the Deskside Onyx by 1000% (i.e., tenfold), thus supporting higher resolution eye tracking data (both in the temporal and spatial domains). Moreover, efficiency of the service has

been increased, so that the required processor usage is now less than 5% of total processor usage (down from more than 50%).

Currently the service provides basic eye tracking information. Our ongoing efforts will further develop the service to include an independent calibration routine, real-time application-based calibration capabilities, and a larger set of high-level, abstract data services (e.g., identification of fixations and saccades). All these activities are pre-requisite as infrastructure to our major NAVCIITI tasks, below.

Task 1: Integrate eye tracker into Dragon software on multiple hardware platforms

We have been continually working to integrate the eye tracking software into Dragon on multiple platforms. To support this goal, we have been working with both hardware and software components.

Early in January, we received our upgraded eye tracking hardware from ISCAN, Inc. The new eye trackers include more robust infrared filters and camera mounting hardware. We have successfully installed the new eye tracking systems and have performed a quality-assurance-check of the upgraded ISCAN eye tracking systems, as well as software tests to ensure that previously developed software was indeed compatible with the new hardware. The new eye tracking systems are currently being used by our NAVCIITI eye tracking software development team with great enthusiasm and at a promising pace.

We have also been working to install the battery backup system for the IWB Deskside Onyx. Final installation of the battery backup is contingent upon installation of a 220-volt power line. We are currently working with the University's physical plant to have the 220-volt line installed in our lab. We expect this task to be complete during the next quarter. At that time we will also install the accompanying battery backup software on our Deskside Onyx to ensure all software development efforts can be reliably stored and "backed up." Again, this is a critical aspect of our infrastructure work upon which our major contributions will be built.

On the software front, we have been refining Dragon's interaction techniques by performing both expert-heuristic evaluation of the user interaction as well as informal pilot testing with users. We have amortized two trips to the Naval Research Lab (NRL) during the quarter such that a portion of the time spent at NRL included evaluation and continued improvement of the Dragon interface, so that it will be even more usable, robust, and efficient for use in our NAVCIITI work.

The Naval Research Lab's Dragon application has been installed and is constantly being upgraded on the IWB. Dragon has not yet been installed on the CAVE, and installation will not be possible until the VT-CAVE is re-constructed (expected during the next quarter).

Task 2: Explore use of eye tracking for data collection for usability analyses

Last quarter, we started designing a "usability evaluator view" that allows evaluators to view the user's scene augmented with eye tracking analysis in real-time on a separate desktop, IWB, ImmersaDesk, or CAVE environment. Software development of the usability evaluator's view was on hold for much of the quarter due to the unavailability of the eye tracking equipment (while it was being upgraded, as discussed above). However, since the return of the eye tracking equipment, substantial software development progress has been made and we expect to have a prototype available during the next quarter.

We expect these features will be extremely valuable in the coming months as we begin to scope the evaluation of DSL/KnowledgeSpace components (currently slated for NAVCIITI year 3).

To support usability analysis, we have extended the user scenarios previously developed for formative evaluation of Dragon to support more thorough and statistically sound, summative usability evaluation of the Dragon application and our emerging user interaction techniques. The scenarios focus on navigation within a terrain-based virtual environment.

Task 3: Explore use of eye tracking as multi-modal interaction technique in VEs

With the return of our eye tracking equipment, we have resumed software development of eye tracking support services as well as multi-modal interaction techniques. As mentioned above, we have refined the user interaction of NRL's Dragon to the point where meaningful multi-modal interaction can be examined without interference from confounding software bugs, hardware bugs, and general inefficiencies.

During this quarter, a research plan for systematically investigating multi-modal interaction techniques has emerged. We will first focus on object selection, since it is a commonly performed task in VEs, and complements the mature navigation established in Dragon. Under our research plan, we will first examine and develop several uni-modal user interaction techniques for object selection. These techniques will then be combined (both logically and in the software) to produce various multi-modal interaction techniques for object selection. A usability evaluation (involving users) will then be performed to examine to what extent multi-modal interaction techniques are more efficient or intuitive than uni-modal techniques.

Significance of Quarterly Effort (with respect to NAVCIITI goals)

Integration of our eye tracking services with DIVERSE is an important step in meeting our NAVCIITI goals. As more NAVCIITI research initiatives are brought together, the ability to evaluate user performance of these systems will be paramount. It is expected that the DIVERSE API will be an integral part of the NAVCIITI integration effort, and as such, our eye tracking component will be well-positioned to provide both innovative user interaction as well as usability and user-task-performance-based insight.

As NAVCIITI begins to examine the creation of a KnowledgeSpace, clearly extensive usability evaluation of various user interfaces will be required. For example, information viewed in a commander's KnowledgeSpace needs to be different than that viewed by the individual combatant. Condensing vast expanses of data and information into timely, useful, and readily digestible knowledge is indeed the essence of human-computer interaction (HCI). As requirements for the DSL/KnowledgeSpace emerge, user-centered evaluation and assessment via eye tracking and other techniques will be critical. We expect our current efforts to transition seamlessly into these forthcoming initiatives.

Activities for the Quarter

- November 2000 – Upgraded Dragon application to reflect latest NRL improvements
- December 2000 – DIVERSE eye tracking service ready for incorporation into next major release of DIVERSE
- December – January 2000 – Visits (one each month) to Naval Research Lab in Washington DC to performing pilot testing of Dragon user interaction
- January 2000 – Improved eye tracking system returned to Virginia Tech
- January 2000 – Upgraded Dragon application to reflect latest NRL improvements
- January 2000 – Extended Dragon user scenarios to support evaluation

Plans for next quarter

Plans for the following quarter include:

- Further develop the eye tracking DIVERSE service to include
  - a calibration routine,
  - hooks for real-time calibration capabilities, and,
  - a larger set of high-level, abstract data services (e.g., fixations and saccades).
- Finalize installation of 220-volt line.
- Install dragon on CAVE (assumes that VT-CAVE will be installed and running).
- Produce a prototype of the evaluator's view running on IWB.
- Produce a prototype of uni-modal user interaction technique for object selection using eye tracking system running on IWB.

Task 2.3.1

Activity Report: Center for Human-Computer Interaction
Navy Collaborative Integrated Information Technology Initiative (NAVCIITI)

9 February 2001

(1) Background on what is the project about. A short description will do.

The responsibility for the Center for Human-Computer Interaction is collaborative software and computer-supported collaborative work (CSCW). We are investigating the development and application of collaborative multimedia conferencing software for education and other groupwork activities. We are extending our own collaborative software framework integrating further tools and capabilities.

(2) Discuss planned effort for the quarter.

Our planned effort for the November-December-January quarter included continuing work on Tasks 2.3.1.3 (collaborative notebook shell), 2.3.1.4 (conferencing components), and 2.3.1.5 (collaborative presentations).

(3) Provide a narrative on what has been done-discoveries, accomplishments, test results,

During November, December, and January the Center for Human-Computer Interaction extended on-going work on software support for collaborative multimedia conferencing, shared notebooks, whiteboard, chat, and visualizations.

Work done on Task 2.3.1.3 during the quarter has included design of a replicated synchronous text editor based on CORK (Content Object Replication Kit), a toolkit developed by the Center for building collaborative software. This editor will serve as a central component of the collaborative notebook, and will also allow us to more easily introduce synchronous editing capabilities into other collaborative tools (e.g., for synchronous editing of individual items in a schedule or outline).

Ongoing work on conferencing tools (under Task 2.3.1.4) has resulted in a preliminary implementation of tools to support audio capture, transmission, and output using platform-neutral Java components. This implementation is currently minimal, providing no call control, limited encoding options, and only a minimal user interface. It is, however, a promising start towards construction of components that address the limitations of current conferencing systems, as identified in the analysis effort during the previous quarter. Specifically, the platform-neutral implementation should allow execution on and interoperability between multiple platforms; the component-based, modular nature of the system should simplify integration into other tools;
and the open nature of the system should allow us to design mechanisms for firewall negotiation and multipoint conferencing.

Work done on Task 2.3.1.5 has included design of a replicated object to support synchronous and asynchronous sharing of presentation data, along with an initial set of user interface components to upload and access shared presentations. In addition to refining the user interface (e.g., to support annotation), ongoing work on this task includes further investigation of tools to support accessing presentation data in third-party formats such as Microsoft PowerPoint files.

(4) Why is it important in accomplishing the objective.

Through the course of the NAVCIITI program, it has become ever clearer that collaboration and collaborative software are key elements of the program, and of its possible continuations. We believe we can play a far greater role in this program as it develops during the balance of the program.

The sensor-to-shooter scenario highlights the tactical importance of sharing and coordinating continuous streams of input in real time. We are focusing on the strategic collaborations that surround such data and make them useful, such as jointly interpreting sensor data, integrating such data with intelligence reports, messaging and negotiation, collaborative annotation (for example, of shared maps), planning and reporting. Many of these collaborations involve shared viewing and editing complex data objects in real time. Most also require transparent integration of synchronous and asynchronous collaboration.

As a member of the National Research Council's Committee on Human Factors, Carroll is now organizing an invitational conference for government agencies to describe their current and future problems with respect to collaborative software, and specifically for supporting remote coordination of work. This is expected to lead to the formation of a National Research Council study panel within the next year.

(5) Activities during the quarter: presentations, publications, workshops, visits, attending / sponsoring conferences, symposia.

Conference papers and presentations in November, December, and January

Carroll, J.M. Scenarios and Scenario-Based Design. Invited Plenary Address for OZ-CHI'2000. Sydney, Australia, December 4, 2000.

Carroll, J.M. 2000. Making use: Scenarios and scenario-based design. In C. Paris, N. Ozkan, S. Howard & S. Lu (Eds.), *Interfacing Reality in the New Millennium, Proceedings of OzCHI 2000.* (Sydney, December 4-8). Canberra: Ergonomic Society of Australia, pages 35-48.

Rosson, M.B. & Carroll, J.M. 2000. Nonfunctional requirements in scenario-based development. In C. Paris, N. Ozkan, S. Howard & S. Lu (Eds.), *Interfacing Reality in the*

New Millennium, Proceedings of OzCHI 2000. (Sydney, December 4-8). Canberra: Ergonomic Society of Australia, pages 232-239.

Carroll, J.M. & Rosson, M.B. 2001, in press. Better home shopping or new democracy? Evaluating community network outcomes. Proceedings of CHI 2001: Conference on Human Factors of Computing Systems. (Seattle, WA; date). New York: ACM, pages xx-xx.

Journal and handbook papers that appeared in November, December, or January, or are now in press

Carroll, J.M. 2000. Scenario-based design. In *McGraw-Hill Yearbook of Science and Technology 2001.* New York: McGraw-Hill, pp. 337-339.

Carroll, J.M. 2000. Introduction to the special issue on human-computer interaction in the new millennium. *ACM Transactions on Computer-Human Interaction, 7/1,* 1-2.

Carroll, J.M. 2000. Introduction to part II of the special issue on human-computer interaction in the new millennium. *ACM Transactions on Computer-Human Interaction, 7/2,* 39-40.

Carroll, J.M. (Ed.) 2000. Preface to collection of Research Alerts (from the double special issue of *ACM ToCHI* on human-computer interaction in the new millennium). *ACM interactions, VII.6,* November/December, 9 (Research Alerts appear on pp. 9-17).

Carroll, J.M. 2000. Scenario-based design. In Waldemar Karwowski. (Ed.), *International Encyclopedia of Ergonomics and Human Factors.* London: Taylor & Francis, pages xx-xx.

Carroll, J.M. In press. Scenario-based design: A brief history and rationale. In C. Eastman, M. McCracken & W. Newstetter (eds.), *Knowing and Learning to Design: Cognitive Perspectives in Design Education.* Amsterdam: Elsevier.

Carroll, J.M., Rosson, M.B., Isenhour, P.L., Ganoe, C.H., Dunlap, D., Fogarty, J., Schafer, W., & Van Metre, C. In press. Designing our town: MOOsburg. *International Journal of Human-Computer Studies*

Carroll, J.M., Rosson, M.B., Isenhour, P.L., Van Metre, C., Schafer, W.A. & Ganoe, C.H. In press. MOOsburg: Multi-user domain support for a community network. *Internet Research.*

Isenhour, P.L., Rosson, M.B. & Carroll, J.M. in press. Supporting interactive collaboration on the Web with CORK. *Interacting with Computers.*

Rosson, M.B. & Carroll, J.M. 2001, in press. Scenarios, objects, and points-of-view in user interface design. To appear in M. van Harmelon (Ed.), *Object Modelling and User Interface Design*. London: Addison Wesley Longmans.

Meetings and related proposal submissions

John Carroll, Dan Dunlap, Philip Isenhour, Dennis Neale and Mary Beth Rosson met with John Rice and Daniel Gilbert of NAVSEA Center for Experimentation and Collaborative Exchange on November 14 in Blacksburg. The purpose of the meeting was to demonstrate the current state of our software development work, and to discuss the methodology we have developed for evaluating the use of collaborative software.

John Carroll, Philip Isenhour, and Mary Beth Rosson met with Tim Pickering of the Virginia Tech Office of Interdisciplinary Programs on January 17. Pickering requested the meeting to discuss disclosure of intellectual property embodied in Center-developed software. Discussions focused on the process for releasing CORK and MOOsburg under an open-source license.

Meeting with Richard Hastings, Associate Director of the VT Alumni Association, to discuss a presentation on collaborative software to the Old Guard Alumni in May.

Submitted two related proposals to NSF to develop collaborative software techniques and applications:

ROLE: The School as a Knowing Organization — Knowledge Management as a Strategy for Continuous Teacher Development
Proposal sent to NSF ROLE program December 1, $710,223 over 3 years
PIs: John M. Carroll, Mary Beth Rosson, Daniel Dunlap, Fred Morton, Robert McCracken, Stephen Kerr, Chun Wei Choo

ITR/PE(IIS): Activity Awareness in Computer-Supported Collaboration
Proposal sent to NSF-ITR program January 24, $499,834 over 2 years
PIs: John M. Carroll, Daniel R. Dunlap, Philip L. Isenhour, D. Scott McCrickard, Dennis C. Neale, Mary Beth Rosson

(6) Plans for next quarter.

Work done by the Center during the February 2001 - April 2001 quarter will include further development of the collaborative authoring tools (Task 2.3.1.3), design of encoding mechanisms and interoperable protocol support for the audio conferencing components (Task 2.3.1.4), and user interface refinement of the collaborative presentation tools (Task 2.3.1.5).

We are also developing plans for Year 3 of the NAVCIITI program. During this year, we would like to develop a demonstration use case for some of the functionality envisioned for the Digital Ship. This activity would help us generalize the collaborative software we

have already developed, and would make more concrete the Digital Ship concept.

(7) Issues if any.

We want to further develop the relationship we have with the NAVSEA Center for Experimentation and Collaborative Exchange. This seems like a valuable Navy contact for us and for the NAVCIITI program more generally.

NAVCIITI Task 2.3.2 & 2.3.3: Digital Ships™ Laboratory

Dr. A. R. Habayeb
NAVCIITI Program Manager
Research and Graduate Studies
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061
Phone: 540-231-4353, Fax 540-231-9805, E-mail: ahabayeb@vt.edu

Introduction

This quarterly progress report describes the objectives, proposed research, current status, and planned efforts for Tasks 2.3.2 and 2.3.3. The goal of Task 2.3.2 is to define the command and control metrics and relate them to the network centric interacting grids. Task 2.3.3 goal is to design, implement and optimize the ship information management system. The Digital Ships Laboratory (DSL) initiative will establish a digital test bed to develop, demonstrate, evaluate, and test advanced shipboard command & control technology. The test bed will be used to integrate intra-ship, and inter-ship Information and provide a network centric focus for the NAVCIITI technology thrusts. A three-tier architecture composed of client, enterprise logic middleware applications, and database tiers will provide a flexible framework for NAVCIITI. The DSL efforts will capitalize and supplement the other NAVCIITI tasks by providing a platform for integrating the research technologies under development.

Objectives

The command and control metrics are implied and assumed ideal in the laws of combat. The formulation of the laws of combat will be examined to derive the command and control metrics**. Simulation of combat environment will be used to establish these metrics.** The DSL provides the mechanism to implement and optimize the ship information management system. DSL is the test bed to integrate command and control technologies. The integration initiative requires sufficient hardware and software for implementation of a flexible, reconfigurable three-tier architecture and the underlying network. The Digital Ships™ infrastructure will be designed with sufficient capacity and flexibility to provide the capability for integration and evaluation of the operational utility of NAVCIITI technologies in a system context. The CAVE facility will be driven by the DSL to provide realism for the simulation.

Efforts for the Quarter and Accomplishments

Efforts for this quarter include progress in three areas: (1) installation of the hardware and software that compose the test platform; (2) definition of documentation standards and approach for requirements gathering; (3) high level decomposition of the command and control system components. Initiated a modeling effort to formulate various scenarios using UML and Use cases.

Accomplishments

The basic hardware for the test platform has been received, configured, and installed in the Digital Ships Laboratory in room 2210, Torgersen Hall. The hardware consists of eight Intel based workstations, one Sun Solaris workstation, and six servers (four rack mounted). Three of the Intel workstations are of sufficient capacity to be used as servers as required. Additionally two 24 port CISCO managed hubs have been included to allow for flexibility in network configuration and support of network traffic monitoring and measurement. This hardware was selected to provide a flexible, modifiable platform to support integration NAVCIITI efforts.

The DSL has been configured to support the Windows 98x/NT/2000, Solaris, and Linux operating systems. The workstations currently support dual boot capabilities for the Windows and Linux operating systems. This capability will be extended to the servers as required.

A thorough understanding of requirements is necessary to provide a basis for design and development of DSL three-tier prototype software. The DSL requirements gathering process will be based on the Unified Modeling Language (UML). The UML defines a Use Case based requirements gathering activity that brings together the functions the system will need to perform. We have documented a requirements definition process that includes a standardized iterative process, a Use Case template, diagram, a business rule table, and other supporting documents. These documents will be incorporated into a working specification. This specification is expected to be dynamic during the life of the project, changing to continuously document the current understanding of system as the research integration effort matures. The Use Case based requirements specification provides the NAVCIITI research team a centralized, standardized means to document and communicate the requirements that each effort will satisfy. As the DSL integration effort matures the requirements documents will serve as the basis for concentration of research efforts, integration of those efforts, and for testing of research prototypes as they are integrated into the DSL.

Three high level system components representing the Sensor-to-Shooter loop have been identified as the Sensor Grid, The Shooter Grid, and the Command and Control grid. Preliminary functional decomposition is ongoing. Diagram 1 depicts how these components may be organized within a collaborative working environment. This working diagram is being used as a basis for documenting functional components identified during walkthroughs of use scenarios, which describe the coordination of the battle space environment. High-level use cases have been documented for each for each of the grids. As the understanding the interactions between the components mature, they will be incorporated in the Use Case specification.

The DynBench, Dynamic Real Time Benchmark Suite and Environment applications being used by the Task 4.1 Scheduling and Resource Mangement team has been installed at DSL. The DSL team is currently exercising the capabilities of the applications and gaining an understanding of the research. Because the software code being used in Task 4.1 models certain functionality identified in the Sensor, Shooter, and Command and Control grids it may be possible to modify these applications for use in the DSL. The effort is ongoing. Efficient resource management is critical to the DSL system concept.

Discussions are also ongoing with the Task 2.2 Visualization and HCI research team. This includes possible requirements for the man-machine interface at individual combat workstations and visualization of the battlespace. Approaches for development of prototype interface software and for evaluating the usability interfaces have been discussed. Discussion has included the needs of the individual sailor under stress conditions and the functions and data that may need to be shared across the various grids. This communication has provided both teams an understanding of the goals of the DSL and the HCI efforts of the NAVCIITI project.

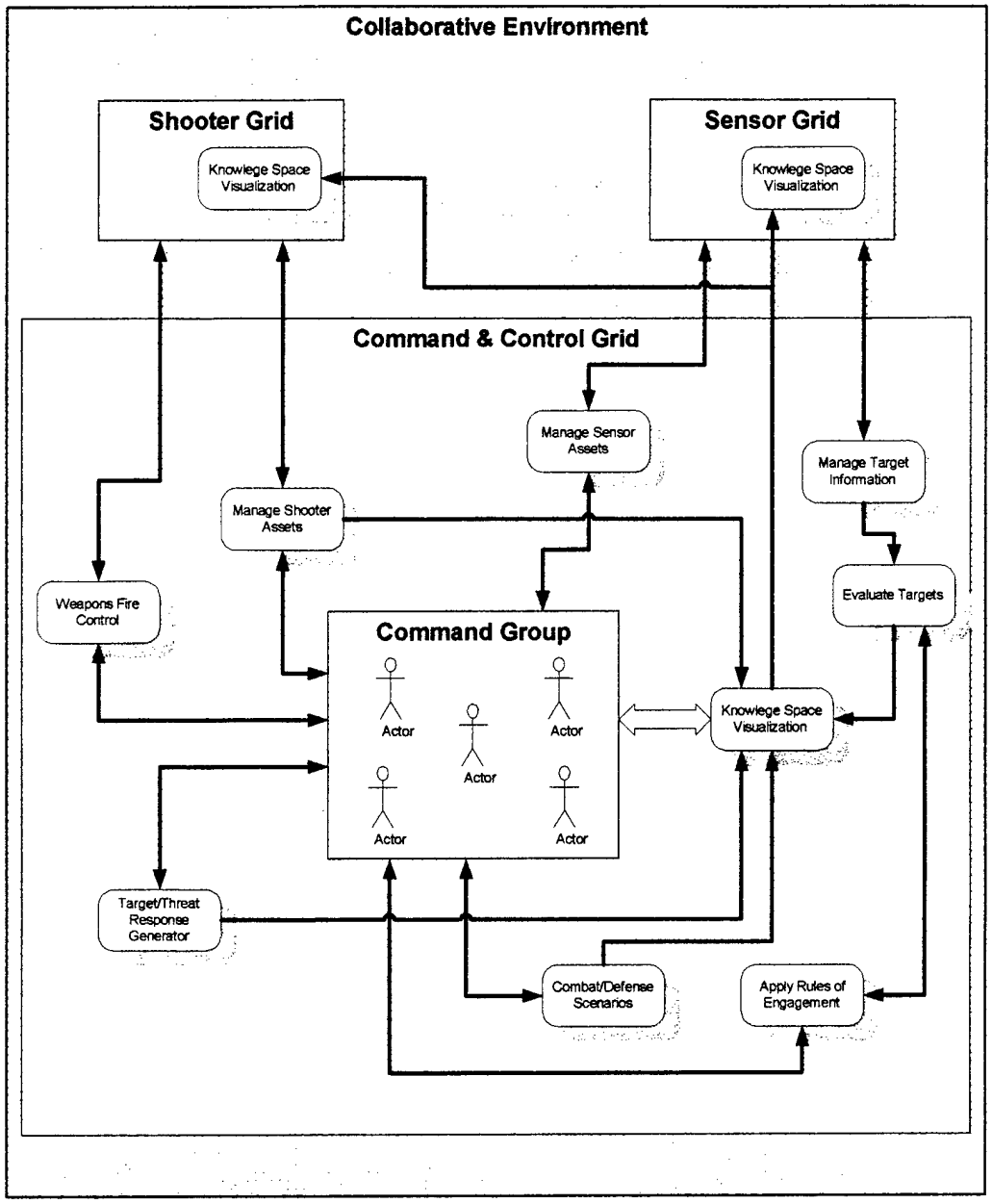


Diagram 1. Working High Level System Component Diagram

Importance of objectives

The realization of the goals of the digital ship effort will result in a cost-effective approach for gaining insights into the many vexing operational challenges facing the fleet. The DSL will provide a platform that will allow advanced command and control technologies to be developed, demonstrated, and tested. The

DSL will serve as a halfway house between NAVCIITI command and control technologies and the fleet. It will serve as a comprehensive testing tool for technology such as, models of antennas, wireless reconfigurable radios, interoperability, and software. The mature system will provide the infrastructure to integrate and evaluate the operational utility of technology in a system context. Also realized will be the ability to perform experiments/runs at a very low cost compared to building physical hardware. NAVCIITI researchers and sponsors will gain further insight into the value of new and brave ideas or concepts by providing a means to examine the interfaces to other technologies and the value of individual research efforts and how they fit with and can be enhanced by other efforts. Also, the information system and components developed and tested in the DSL can provide a training platform for Navy personnel.

Activities

The DSL concept and the Sensor-to-shooter grids were presented to personnel from NAVSEA NSWC Dam Neck and New Port News Shipbuilding on Dec 14,2000.

Plans for Next Quarter.

During the next quarter the DSL team will further develop the projects described above. Further decompositon and documentation of the requirements for the Sensor, Shooter, and Command and Control grids will be performed. The goal is to refine requirements for specific client simulators, middleware components, and data storage to allow development of preliminary prototype applications for the three-tiers. These application will be modified during the project life-cycle to allow integration the the NAVCIITI research efforts. The inherent modifiability of the UML, object oriented development approach being used for the DSL allows this iterative, incremental development. Appropriate network measurement capabilities will incorporated to establish baseline metrics as needs are identified.
Interaction with the NAVCIITI team is ongoing to create collaboration and sharing of resources.

2.33 Modeling Ship on fire

Background

The purpose of this effort is to provide a simulation of ship operations and response when part of the ship is exposed to a fire. Certainly fires can affect the personnel, equipment, and structures on a Navy ship and the degree to which each of these is affected depends not only on the fire itself, but how the ship's systems and personnel respond to the situation (dictated by command and control procedures). To establish as complete a fire simulation as possible, several distinct tasks are being integrated. First, the evolution of the conditions on-board a ship due to the presence of a fire is being predicted. Second, the influence of these conditions on the ship's personnel, equipment, and structures is being modeled. Ultimately, the information from these models will be linked with the Navy's command and control methodologies and the entire process will be simulated in real-time on a virtual network.

Objectives

One of the main objectives for this quarter was to develop computer models to complement existing fire models through the prediction of fire-effects on materials or structures on board a ship – an important consideration in the ship-on-fire simulations. The second objective was to design and construct the special experimental equipment required to validate the fire and structural response models and to conduct preliminary experiments.

Accomplishments

During the previous quarter (Y2Q2), efforts focused mainly on setting up and learning the basic principles of the FDS fire model. During this quarter, steps were taken to develop a greater familiarity and more complete working knowledge of the program. The capability to predict and visualize data on air temperature, gas concentrations (both toxic and non-toxic), and smoke density has now been firmly established. This information, when compared with habitability or operability standards, will directly reveal the effects of fire on a ship's personnel and electronic equipment – two important factors affecting command and control of a ship on fire.

Since it is clear that fire on a ship not only affects personnel and equipment, but also on-board structures, attention was turned this quarter to studying the response of mechanically-loaded structures to fire exposure. To this end, several thermo-mechanical response models have been developed. These include a heat transfer model to compute the temperature distribution within a material sample based on the radiative and convective heat fluxes as well as elevated ambient temperatures predicted by the FDS model. A temperature-property model was then introduced to compute the material property distribution within the structure as a function of time. Finally, a mechanical response model is used to compute the time-dependent deformation of a beam (with non-uniform stiffness) in bending exposed to fire. Additional models will be added to these in the future to predict when the combination of mechanical loading and material property degradation due to fire exposure result in structural failure.

To enable the experimental study of structural response to fire and to validate the model discussed above, a special fire-testing chamber was designed and constructed during this quarter. The 5 ft^3 chamber is completely enclosed (by steel walls on three sides and on top and by a glass observation window on the front) except for venhilation ports in order to sustain combustion within the chamber. A diffusion burner (fueled by propane gas) can be placed at any location on the floor of the chamber and provides controlled fires ranging in size from 0.1 kW to 0.3 kW. Thermocouples are used to monitor the air temperatures within the chamber over time as the fire burns and fills the chamber with hot gases. Material samples can be placed inside the burning chamber in order to study the combined effects of mechanical loading and fire exposure. The chamber was carefully designed such that the magnitude of the load as well as the location of the sample relative to the fire can be varied over a wide range.

Several preliminary tests have already been performed using the new experimental equipment. The air temperature measurements have been compared with predictions from the FDS model for several different fire sizes and good repeatability and agreement has been observed. Tests of materials subjected to combined fire and mechanical loading have also been conducted and the times-to-failure for structures over a range of fire and loading conditions have been measured. Though the experimentally-observed beam deflections are qualitatively similar to the deformations predicted by the thermo-mechanical models discussed above, a quantitative comparison has not yet been made.

The work described in this report is receiving attention from the scientific community as evidenced by the acceptance of three abstracts for oral presentation at upcoming conferences. The first presentation will occur in late February at the Adhesion Society Meeting in Williamsburg, VA and will focus on the response and failure of adhesively bonded structures exposed to fire. Later in the year more extensive fire modeling and experimental work will be presented at the American Society of Composites meeting in Blacksburg, VA and at the International Conference on Fracture meeting in Honolulu, HI.

Future Work

Over the next quarter, efforts will be focused on experimentally measuring the response of structures exposed to simultaneous fire and mechanical loading. The results will be used to help refine the thermo-mechanical response models developed during this quarter. The thrmo-mechanical response models will then be integrated with the existing FDS fire model in such a way that the effects of fire on a ship's structures, personnel, and equipment can be predicted simultaneously (with one set of models). Ultimately, these models will be linked with command and control models provided by the Navy and the Digital Ship network will be used to provide a simulation of the operation of a ship on fire in real-time.

Task 2.4

Mechanically Flexible Multifunctional Display Materials and Devices

Liangmin Zhang, Roger Duncan, Fajian Zhang, Tingying Zeng and Richard O. Claus
Fiber & Electro-Optics Research Center
Virginia Tech

1. Background

The objective of this work is to investigate methods for the formation of multifunctional display materials and devices for use in the presentation of data for Digital Ship visualization applications. This quarter we have investigated in particular the light detection capability of self-assembled film materials that may be integrated with display elements to allow integrated laser pointing detection, response and display by a single device.

2. Planned Effort for Quarter

Our plan for this quarter has been to demonstrate the light detecting capability of self-assembled protein-based thin film materials.

3. Accomplishments

Our accomplishments are summarized in detail in this section. In general summary, the light detection behavior of thin films in which bacteriorhodopsin (bR) light detecting protein molecules are embedded in a polyvinyl alcohol matrix has been investigated by using both pulsed laser excitation and regular light illumination. This has been done to simulate the exposure of the top surface of a display element to laser pointer excitation in a Digital Ship interactive display application.

Excellent response times as short as milliseconds, strong photocurrents as great as 120 $\mu A/cm^2$, and high photovoltages as large as 3.8 V have been obtained. Two theoretical models have been developed and used to extract several physical parameters and fit the experimental results, respectively. Some important intrinsic parameters have been obtained. Theoretical results show

that excited protons can construct an internal electric field of 10^2 V/m under an illumination intensity of 1.5 W/cm^2 when no external voltage is applied, and the lifetime of excited protons is ~10^4s. Theoretical results also indicate that the average displacement of the excited protons is on the order of several tens of microns. Other curve fits show that photocurrent and photovoltage increase linearly with external field, but increase exponentially with flash period and flash power. These theoretical models and results can be extended to other kinds of optically active polymeric materials, and have important implications in areas other than multifunctional display elements.

3.1 Review of Optical Detection Materials and Processes

The supramolecular assembly of optically active materials into thin film architectures is a topic of current theoretical and experimental interest and of significant practical importance in photoelectric devices. Bacteriorhodopsin (bR) is the light-harvesting protein of the purple membrane (PM) of the halophilic microorganism Halobacterium halobrium [1-5]. Uniformly oriented bR molecules in PM perform unidirectional pumping of protons from the cytoplasm to the extracellular space during the photochemical cycle of bR [6-9]. Although the exact mechanism of the proton channel is still unclear and is currently explained using various models [5,10-11] generally, the Schiff base, hydrogen bond network (formed by several water molecules along the route of proton translocation), and amino acid residues Asp85 and Asp96 (where Asp stands for aspartic acid) are involved in the formation of the proton channel. Fig. 1 shows the generally accepted scheme for the bR photochemical cycle [5-6, 10-11, 12-16]. The long dashed-lines in this cycle indicate photon-driven processes and the solid-lines correspond to thermal transformations. The short dashed-lines show the release and uptake processes of protons. The numbers in brackets indicate the absorption maximum wavelengths in units of nm.

In the L to M reaction, the protons are released to the extracellular (EC) side. The deprotonated retinal chromophore then takes up a proton from Asp96, which subsequently regains a proton from the cytoplasmic (CP) side in the M to N reaction.

Unlike typical biological materials, PM has high thermal and photochemical stability. bR can retain its natural structure and function to temperatures as high as 140° C in the dry state, and for pH ranging from 3 to 10 in solution [6, 17]. The proton-pumping photochemical cycle of bR can be repeatedly cycled more than 10^6 times [5,10]. This unusual stability coupled with its unique photochemical and photophysical properties have made bR the most promising biological material for photonic device applications.

Although the description given above is an oversimplification of the complicated function of bR it clearly demonstrates that the photochemical cycle of bR consists of states with different absorption maxima, the shifts of which are strongest for the deprotonation and reprotonation step of the Schiff base. With its photochemical cycle, bR represents a reversible photochromic system in addition to its charge translocation property. Since the M-state exhibits the strongest spectral shift and the highest population of all intermediates under steady-state light conditions, one must try to extend the lifetime of the M-state to obtain the strongest possible photocurrent signal. Therefore, the lifetime of the M-state has received considerable interest in nearly all optical and electro-optical applications of bR. So far, many approaches have been used in attempts to control the lifetime, including physical (applying an external voltage to the materials [18,19]) and chemical (changing the pH value of the solutions and adding sodium azide to the electrolyte [6,9,13]) methods, as well as conventional mutagenesis [20] and genetic

engineering [6, 7, 12, 21]. Lifetimes as long as several tens of seconds and photocurrents as great as several hundred nanoamperes per square centimeter have been obtained. In this discussion, we report results for bR embeded in a polyvinyl alcohol (PVA) polymer matrix by the sol-gel technique. The lifetime of the M-state is extended dramatically to as long as 10^4 s, and a photocurrent amplitude as great as 120 $\mu A/cm^2$ is obtained. Moreover, the photocurrent and the lifetime as functions of flash power, flash period and external voltage are shown.

3.2 Materials and Setup

PVA (Aldrich, 99% with Mw=89,000-98,000) was used as the matrix material, and purple membrane (PM) powder from Halobacterium Halobium (Sigma, 75% bacteriorhodopsin) was employed as the bR source. To synthesize high optoelectric response oriented bR/PVA composite films, a biomimetic templating sol-gel "soft" chemical approach was used. Thus, the organic molecule 3-(1-pyridinio) propane sulfonate (PPS), one of the agents chosen to improve the recovery of membrane proteins, was used as the template to induce the bR molecules to be self-assembled during processing. Therefore, the bR molecules encapsulated in PVA matrix are self-assembled into an oriented nanostructured film with the PPS. The preparation involves three procedures. First, 0.05mL 4.7% PPS (Fluka chemika) aqueous solution was mixed with 0.2mL 15% PVA aqueous solution. Then, a small amount of sodium azide Na_3N was added into the mixture as an anti-corrosion agent, the pH of the mixture is 7.0. Second, a PPS modified bR/PVA sol was prepared by adding 0.2mL bR phosphate buffer solution (pH=7.0) with a bR concentration of 1.88 mg/L to the mixture obtained above, and mixing the sol thoroughly. Third, the bR/PVA sol was coated uniformly onto carefully cleaned ITO-coated

glass by spin coating. The as-prepared bR/PVA composite film was dried in air for four days at room temperature. Films with thicknesses of 101.6 μm were obtained. From the dried film volume, a bR molecule concentration of 1.77×10^{19} molecules/cm^3 in the dry film was calculated.

Fig. 2 depicts the UV-vis absorption spectrum of a bR/PVA thin film with a thickness of 101.6 μm, recorded using a Hitachi-U2010 spectrophotometer. One can see that the peak absorption wavelength is approximately 560 nm. The characteristic absorption band of the bR due to the retinal chromophore is still observed after the films were dried for four days. It indicates that the self-assembly of the bR protein in the PVA polymer matrix modified by PPS does not make the protein denatured. The setup employed to measure the photocurrent generated by the thin film is shown in Fig. 3. In this setup a DC voltage from a high voltage power supply (Bertan 205B-01R) can be applied across the two electrodes, one of which is an ITO coating on a glass substrate and the counter electrode is fabricated using a self-adhesive copper tape (diameter~3mm). A rectangular pulse of a continuous wave laser beam with a diameter of 3 mm from a Coherent Verdi 5 laser oscillating at 532 nm illuminates the sample intermittently. The period of the light flash is controlled by a digital shutter (Newport 845-HP). A reference signal from an external fast photodiode D arranged to detect a small sampled part of the primary beam provides a trigger input for the storage oscilloscope (Tektronix TDS-430). In the presence of photoconduction, current flows in the external circuit, and a voltage drop across a resistor R is measured by the storage oscilloscope.

3.3 Theory and Results

3.3.1 No external voltage applied

Fig. 4 represents a typical photocurrent signal obtained when no external voltage was applied to the sample. The flash power is 105.1 mW and the flash period is 40 s, as shown in the lower part of this figure. The inset shows a curve fit of the data to model the decay process after the flash power is shut off. The negative photocurrent indicates that protons move from the metallic electrode to the ITO-coated substrate in the sample. This observation shows us that the bR molecules are oriented in the films even though no external voltage is applied, and that the EC side faces the ITO-coated substrate because the pumped protons are transported from the CP side to the EC side across the PM while light illuminates the sample. One can see that the photocurrent curve has an approximately rectangular shape. At the plateau shown, the photocurrent reaches its quasi-steady state. To extract some physical parameters from this behavior, let us consider the following model. Assume that the coordinate of the ITO electrode is at the position x = 0 and the counter metallic electrode is at x = D (the thickness of the samples). Because the area of the electrode is sufficiently larger than D, as an approximation, fringing fields may be ignored and the internal electric field vector points in the x direction [22]. When a laser beam irradiates the sample at x = D, the pumping protons move from the CP side to the EC side, or macroscopically from the metallic electrode to the ITO substrate. Maxwell's equation for the amplitude of the electric field in the x direction gives

$$\frac{dE}{dx} = \frac{Q_1}{\varepsilon_0 \varepsilon}\delta(x) + \frac{Q_2}{\varepsilon_0 \varepsilon}\delta(x-D) + \frac{\rho(x)}{\varepsilon_0 \varepsilon}, \qquad (1)$$

where ε_0 and ε represent the dielectric constants of vacuum and the film, respectively. Q_1 and Q_2 are the charges on the ITO electrode and the metallic electrode, respectively, and the $\rho(x)$ is the spatial charge density distribution function of excited protons. By using the charge density per unit area $\sigma(x)$ and integrating equation 1, one can modify equation 1 to obtain

$$E = \frac{Q_1}{\varepsilon_0 \varepsilon} + \frac{Q_2}{\varepsilon_0 \varepsilon} + \int_0^D \frac{\sigma(x)s}{\varepsilon_0 \varepsilon} dx, \qquad (2)$$

where s represent the area of the incident pulsed beam. As an approximation, we assume the proton area density distribution is given by

$$\sigma(x) = \sigma_0 \exp(-kx), \qquad (3)$$

where σ_0 and k are two constants. Then, one can obtain $Q_1=\sigma_0 s$ and $Q_2=\sigma_0 s\exp(-kD)$. The boundary conditions are

$$Q_1 + Q_2 + \int_0^D \sigma(x) s dx = Q_0, \qquad (4)$$

and

$$\int_0^D E dx = V_0, \qquad (5)$$

where Q_0 is total charge in the illuminated volume and V_0 is the open circuit voltage between the two electrodes. In Fig. 4, the flash beam was shut off at 40 s, and the excited photocurrent decreases to zero at 50 s. By integrating the fitted photocurrent curve from 40 s to 50 s, one can obtain the area enclosed by the two dashed lines and the fitted curve, giving the quantity Q_0 = 6.7 nC at 40 s when the flash beam was shut off. By taking ε =1.8 and using the equations above, one can obtain $E=4.2\times10^2$V/m and V_0=42.9 mV. We measured the open circuit voltage

between the two electrodes during the quasi-steady state conditions to be 46±5 mV. This essentially agrees with the theoretical calculations. In addition to these, one can obtain the number of the pumping protons N_t in the illuminated cylinder during the quasi-steady state by using the Q_0 value of 6.7 nC, giving $N_t = 4.2\times10^{10}$. From the concentration of bR molecules in this sample, one can also approximately calculate the number of bR molecules in this illuminated volume to be $N\sim10^{17}$. From this data, one can infer that only $1/10^7$ of the bR molecules in this volume contribute to the photocurrent without an external voltage applied.

3.3.2 Effects of applied external voltage

Figs. 5, 6 and 7 show different kinetic features of photovoltage and photocurrent for different flash periods and 20 V external voltage (corresponding to an electric field of $E = 1.97\times10^5$ V/m) applied across the sample. The flash pulses are also shown in the lower parts of these figures. In Fig. 5, the flash period is 30 s and flash power is 154.1 mW. One can see that the maximum photocurrent density reaches 120 $\mu A/cm^2$ in Fig. 5. In Figs. 6 and 7, the flash power is 129.0 mW and the flash periods are 60 ms and 10 ms, respectively. One can see that these films respond to illumination very fast. Fig. 8 shows the dependence of photovoltage on the flash power. The flash period is 5 s and the external voltage is 20 V. One can see from this figure that the flash power can affect the photovoltage greatly, and a maximum photovoltage as high as 1.72 V is obtained when a flash power P=170.2 mW is used. In all the figures above, the solid lines indicate the theoretically fit results and the symbols show measured data. Before discussing these results in detail, let us review and develop a theoretical model [23-25]. Let us consider a single oriented bR molecule in a homogeneous sample. An absorbed photon excites

a proton to displace a distance d from point 1 to point 2 inside the sample. A voltage V(t) is then created in the external circuit loaded by a capacitance C and the resistance R,

$$V(t) = \frac{ed}{DC}\exp(-t/RC), \qquad (6)$$

where e is the elementary charge, D is the distance between two electrodes, in this case, the thickness of the sample, C is the capacitance of the sample, and R=0.5 MΩ in our measurement setup. In a real experiment, at t = 0, the photons excite N_0 protons. Because the excited protons are taken up by the transition from the M-state to the N state, the number of protons $N_t(t)$ undergoes a simple exponential decay

$$N_t(t) = \frac{N_0}{\tau_1}\exp(-t/\tau_1), \qquad (7)$$

where τ_1 is the lifetime of the protons. Every excited-proton displacement produces a voltage as given in Eq. (6). To obtain the total voltage $V_N(t)$ we have to integrate the V(t) function for all time t'<t, this means the folding of Eqns. (6) and (7) into

$$V_N(t) = \frac{N_0 ed}{DC\tau_1}\int_0^t \exp(-t'/\tau_1).\exp[-(t-t')/RC)dt'$$

$$= \frac{N_0 ed}{D}.\frac{R}{\tau_1 - RC}[\exp(-t/\tau_1) - \exp(-t/RC)]. \qquad (8)$$

In our case, the area of the metallic electrode is ~ 7 mm^2, the capacitance is ~ 50 pF, and RC ~ 25 μs. The lifetime τ_1 of the protons is much longer than RC, so Eq. 8 can be simplified to

$$V_N(t) = \frac{N_0 ed}{D}.\frac{R}{\tau_1}\exp(-t/\tau_1). \qquad (9)$$

After being excited, protons will be taken up and re-trapped in the photochemical cycle. Therefore, the proton velocity v should be a function of time. To describe this process more accurately, let us assume the velocity of the moving protons is

$$v(t) = \frac{d}{\tau_2}\exp(-t/\tau_2), \qquad (10)$$

where τ_2 is the rate constant that describes the proton uptake process. After taking this effect into account and folding Eqs. (9) and (10), we obtain

$$V_N(t) = \frac{N_0 edR}{D(\tau_1 - \tau_2)}[\exp(-t/\tau_1) - \exp(-t/\tau_2)]. \qquad (11)$$

After the flash power is turned off, the photovoltage and photocurrent decrease quickly. To describe this process, two parameters must be introduced into Eq. (11)[25] by

$$V_N(t) = \frac{N_0 edR}{D(\tau_1 - \tau_2)}[A\exp(-t/\tau_1) - B\exp(-t/\tau_2)], \qquad (12)$$

where A and B are two constants. One can see from Figs. 5-8 that the curves fitted by using this model are essentially in agreement with the experimental measurements although the flash period varies from 10 ms to 30 s and the flash power varies from 11mW to 170 mW. Figs. 9 and 10 show the fitted lifetime τ_1 (●) and uptake rate constant τ_2 (■) of excited protons during illumination as functions of flash power and externally electric field. One can see from these figures that the lifetime τ_1 is on the order of several tens of thousands of seconds and the rate constant τ_2 is ~10^4 times smaller than τ_1. Also, one can see that τ_2 is essentially not affected by the flash power and the external field, and τ_1 is significantly dependent on both the power and electric field. One can infer that τ_2 is determined by intrinsic factors of the sample.

In addition to τ_1 and τ_2, one can obtain the average displacement d of the excited protons during illumination from the curve fit by using Eqs. (11) and (12). By integrating the fitted photocurrent curves with respect to time from the time when the flash power is turned off to the time when the photocurrent decreases to zero as in Fig.4, one can obtain the number N_t of protons at the time when the flash power is turned off. The results obtained by integrating are shown in Fig. 11. Symbols indicate the results of integrating the nine fitted curves in Fig. 8 and the solid-lines correspond to the fitted result. The curve fit shows $N_t \sim P^{1.35}$, where P stands for the flash power in units of mW. Comparing the value N_t obtained here with the result in Fig. 4, one can see that N_t is $\sim 10^3$ times larger than that without an applied external voltage. This is because when an external field is applied across the sample, the orientation of the bR molecules in the sample is improved greatly. Fig. 12 shows the fitted average displacement of excited protons as a function of the flash power. Unlike the results obtained for suspensions of electric field–oriented PM reported in Refs. 23 and 24, in our sample, the average displacement is several tens of micrometers instead of several nanometers. A possible explanation is that our samples are solid films and bR molecules are embedded in a PVA matrix that acts as a proton transport species. Thus the protons can easily hop all around inside the sample by means of hydrogen bonds. In contrast, in oriented liquid samples, the protons can only move within a single fragment (size ~ several nanometers) of PM since the individual fragments are isolated by solution. This is one of the reasons why we obtain much stronger photocurrents. It is also worthwhile to note that the average displacement only slightly varies with increasing power. This may indicate that the displacement is mainly determined by the composition of the samples.

Fig. 13 shows photovoltage and photocurrent intensity as a function of the external field. Symbols indicate the measured results and the solid-line corresponds to the fitted result. The flash period and flash power are 10 s and 129.0 mW, respectively. The fitted function is $V(mV) \sim (E_0)^{1.06}$, where V is photovoltage and E_0 is the external field. One can see that the photovoltage increases linearly with the external field for this kind of material. Fig. 14 shows photovoltage and photocurrent as a function of flash period. Again, symbols indicate the measured results and the solid-line corresponds to the fitted result. The external voltage and flash power are 20 V and 129.0 mW, respectively. The fitted function is $V(mV) \sim (\Delta t)^{0.18}$, where Δt is the flash period and in units of ms. From this figure, one can see that the photovoltage increases functionally more slowly with increasing flash period than with external voltage. Fig. 15 shows photovoltage and photocurrent intensity as a function of the flash period. Symbols indicate the measured results and the solid-line corresponds to the fitted result. The flash period and external voltage are 5 s and 20 V, respectively. The fitted function is $V(mV) \sim (P)^{1.66}$, where P is the flash power in units of mW. One can see from this fit that of the parameters studied, the photovoltage and photocurrent increase most quickly with increasing flash power. Therefore, increasing the flash power is most the effective method we have identified to obtain high photovoltage and large photocurrent in self-assembled bR films.

Considering the potential practical applications of such materials in optoelectronic devices, we used a projector (Fiber-Lite 3100, Dolan-Jenner Industries) as a CW visible light source to illuminate the sample. The output light intensity from this projector is 11.1 W/cm^2. No filter was used. For an external voltage of 25 V applied to the sample, and illumination period of 68

s, the kinetics of photovoltage is shown in Fig. 16. Symbols and the solid-line indicate the measured and fitted results, respectively. The obtained maximum photovoltage is 2.1 V, corresponding to a photocurrent intensity of 59μA/cm^2. The fitted results according to Eq. 11 give $\tau_1 = 2.36\times10^6$s and $\tau_2 = 8.79$ s. This result shows that the sample materials can also produce a strong photocurrent upon illumination of by regular CW visible light.

3.4 Conclusions for Quarter Research of Display Detector Device Properties

In summary, versatile optical functions of bR retinal proteins are highly promising for use in photovoltaic devices, optical modulation and information processing. These require the highest possible sensitivity toward light illumination. We have fabricated a novel type of bR/PVA films using a modified sol-gel technique, and have demonstrated their high-performance photovoltaic behavior by using both pulsed laser excitation and regular light illumination methods. In these films, the lifetime of excited protons is extended dramatically to as long as 10^4s. As a result, the corresponding maximum photovoltage and photocurrent intensities reach 3.8 V and 120 μA/cm^2, respectively. Comparing these results with those reported previously, the photocurrent intensity is several tens of times larger. A theoretical model is presented and used to estimate the electric field created internally by the excited protons in the case without an externally applied voltage. It is found that the excited protons for a laser illumination intensity of 1.5 W/cm^2 can produce an electric field on the order of 10^2 V/m when no external voltage is applied. Another model has also been developed and used to extract several intrinsic parameters and fit the kinetics of photovoltage and photocurrent. The experimental results agree with the fitted results. The results indicate that protons may be able to hop all over within

the films and the average displacement is on the order of several tens of micrometers as well as illumination intensity dependent. Theoretical fits also show that the photovoltage and photocurrent increase linearly with applied external field, but increase exponentially with flash period and flash power.

3.5 References

1. K. Schulten and P. Tavan, Nature, 272, 85(1978).
2. R. Shinar, S. Druckmann, and M. Ottolenghi, R. Korenstein, Biophys. J., 19, 1(1977).
3. P. Ormos, Z. Dancshazy, and L. Keszthelyi, Biophys. J., 31, 207(1980).
4. R. R. Birge, T. M. Cooper, A. E. Lawrence, M. B. Masthay, C. Vasilakis, C. F. Zhang, and R. Zidovetzki, J. Am. Chem. Soc., 111, 4063(1989).
5. H. Luecke, H. T. Richter, J. K. Lanyi, Science, 280, 1934(1998).
6. J. A. He, L. Samuelson, L. Li, J. Kumar, and S. K. Tripathy, Adv. Mater.,11, 435(1999).
7. K. Koyama, N. Yamaguchi, and T. Miyasaka, Adv. Mater., 7, 590(1995).
8. L. Song, M. A. El-Sayed, J. K. Lanyi, Science, 261, 891(1993).
9. J. A. He, L. Samuelson, L. Li, J. Kumar, and S. K. Tripathy, J. Phys. Chem. B, 102, 7067(1998).
10. E. P. Peyroula, G. Rummel, J. P. Rosenbusch, E. M. Landau, Science, 277,1676(1997).
11. R. Korenstein, B. Hess, Nature, 184, 270(1977).
12. K. Koyama, N. Yamaguchi, and T. Miyasaka, Science, 265, 762(1994).
13. C. Brauchle, N. Hampp, and D. Oesterhelt, Adv. Mater., 3, 420(1991).
14. H. J. Butt, K. Fendler, E. Bamberg, J. Tittor, and D. Oesterhelt, EMBO J. 8, 1657(1989).

15. B. Robertson and E. P. Lukashev, Biophys. J., 68, 1507(1995).

16. Z. Chen, A. Lewis, H. Takei, and L. Nebenzahl, Appl. Opt., 30, 5188(1991).

17. Y. Shen, C. R. Safinya, K. S. Liang, A. F. Ruppert, K. J. Rothschild, Nature, 366, 48(1993).

18. E. P. Lukashev, E. Vozary, A. A. Kononenko, and A. B. Rubin, Biochim. Biophys. Acta, 592, 258(1980).

19. S. K. Chamorovsky, E. P. Lukashev, A. A. Kononenko and A. B. Rubin, Biochim. Biophys Acta, 725, 403(1983).

20. J. Soppa, D. Oesterhelt, J. Biol. Chem. 264, 13043(1989).

21. T. Mogi, L. J. Stern, T. Marti, B. H. Chao, H. G. Khorana, Proc. Natl. Acad. Sci. USA, 86, 529(1988).

22. A. Hirao, H. Nishizawa, and M. Sugiuchi, Phys. Rev. Lett., 75, 1787(1995).

23. L. Keszthelyi and P. Ormos, FEBS Lett., 109, 189(1980).

24. P. Ormos, L. Reinisch, L. Keszthelyi, Biochim. Biophys. Acta, 722, 471(1983).

25. P. Lauger, R. Benz, G. Stark, E. Bamberg, P. C. Jordan, A. Fahr and V. Brock, Q. Rev. Biophys., 14, 513(1981).

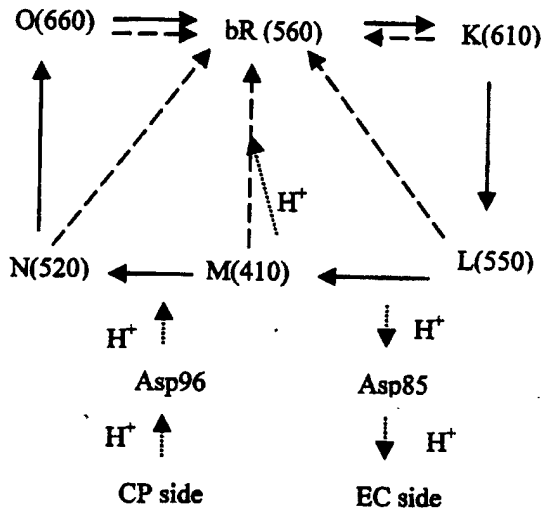


Fig.1. Photochemical cycle of the bR molecule. The long dashed-lines and arrows indicate the photon-driven process. The solid-lines and arrows show the thermal decay paths. The short dashed lines and arrows indicate the release and uptake processes of protons.

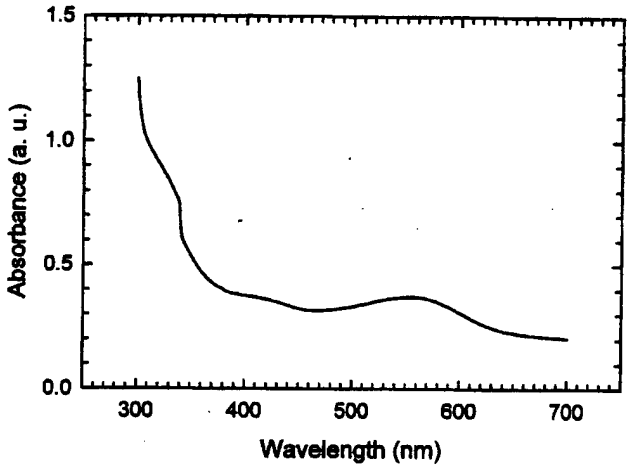


Fig. 2. UV-vis absorption spectrum of the bR/PVA film.

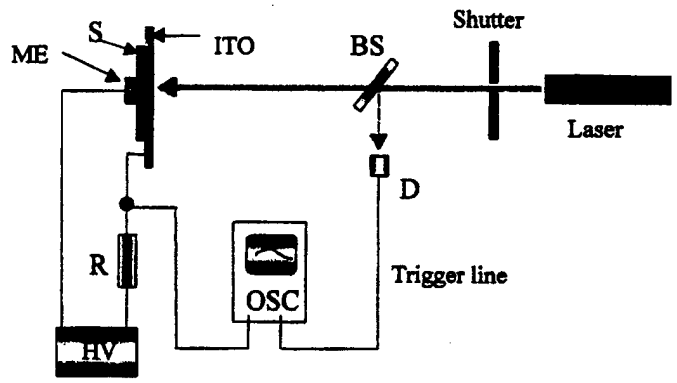


Fig. 3. Experimental setup for the photocurrent measurements.

ITO: indium-tin-oxide coated glass; S: sample; ME: copper electrode; BS: beam splitter; D: photodiode; R: resistor (0.5 MΩ); Shutter: digital shutter; OSC: storage oscilloscope; HV: high voltage power supply; Laser: Coherent Verdi 5 oscillating at 532 nm.

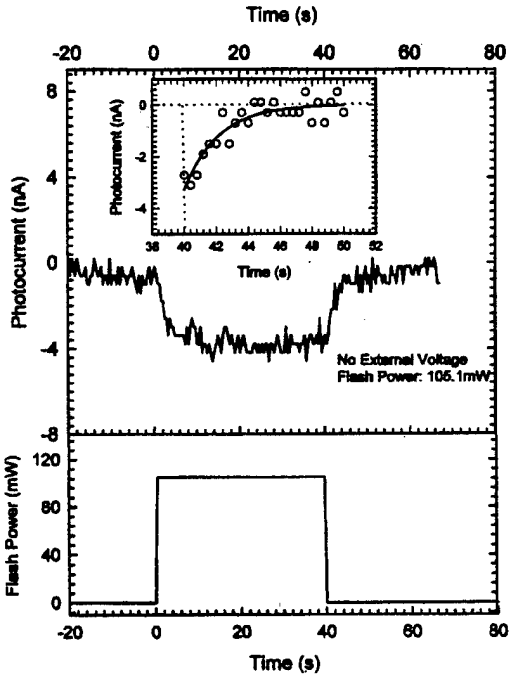


Fig. 4. Photocurrent kinetics without external voltage applied. The inset shows the fitted results for the time period from 40 s when the flash beam was turned off through 50 s when the photocurrent decreases to zero. The lower part shows the shape of the laser pulse controlled by a digital shutter.

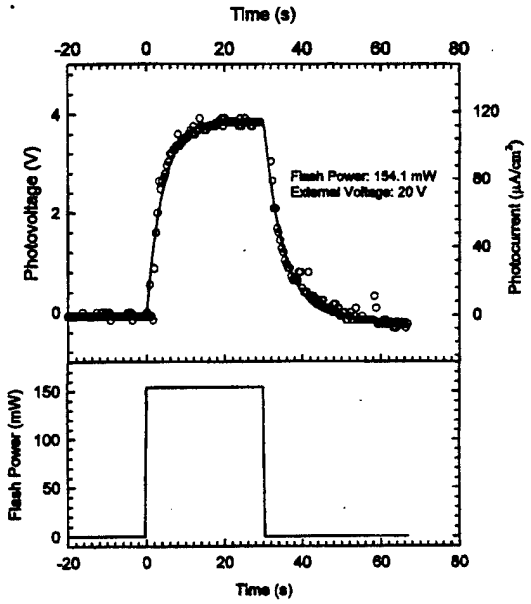


Fig. 5. Photovoltage and photocurrent as a function of time. An external voltage of 20 V is applied (corresponding electric field: 1.97×10^5 V/m) across the sample. The lower part of this figure shows the laser pulse. The pulse power is 154.1 mW. Symbols show the experimental results and the solid-line indicates the fitted result. One can see that the maximum photovoltage and photocurrent intensity reach 3.8 V and 120 $\mu A/cm^2$, respectively, and the theoretical model gives a good description for the process.

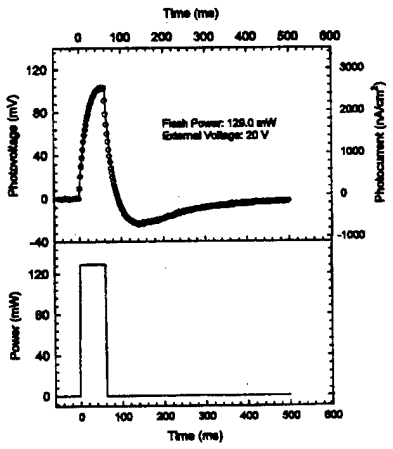


Fig. 6. Photovoltage and photocurrent as a function of time. External voltage is 20 V. Flash period is 60 ms. Symbols show the experimental results and the solid-line indicates the fitted result. The lower part shows the laser pulse.

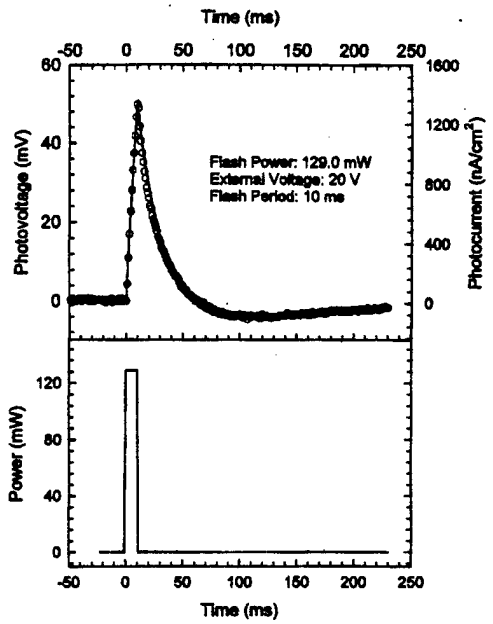


Fig. 7. Photovoltage and photocurrent as a function of time. External voltage is 20 V. Flash period is 10 ms. Symbols show the experimental results and the solid-line indicates the fitted result. The lower part shows the laser pulse. One can see from this figure that the photovoltage has a quite short response time upon illumination.

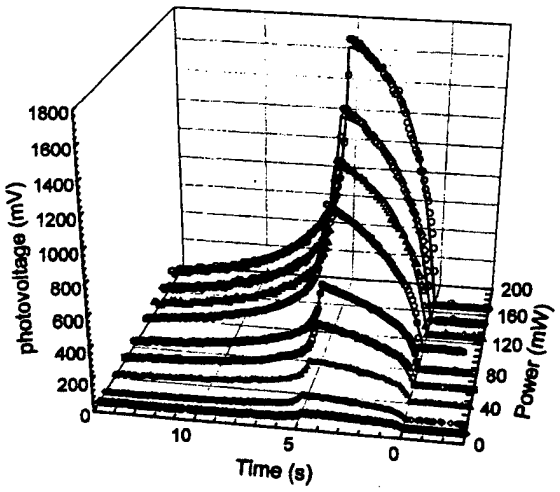


Fig. 8. Kinetic feature dependence on power under the same external voltage of 20 V and flash period of 5 s. Symbols show the experimental results and the solid-lines indicate the fitted results.

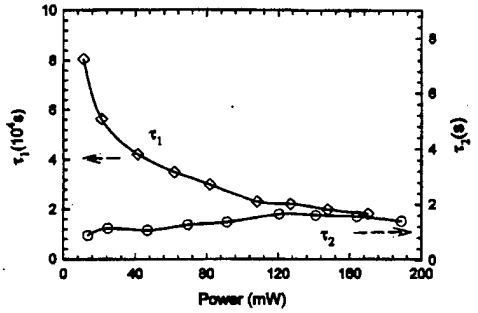


Fig. 9. The fitted lifetime τ_1 (♫) of excited protons and the rate constant τ_2 (♦) according to Eq. 11 as a function of the flash power. The external voltage and flash period are 20 V and 5 s, respectively.

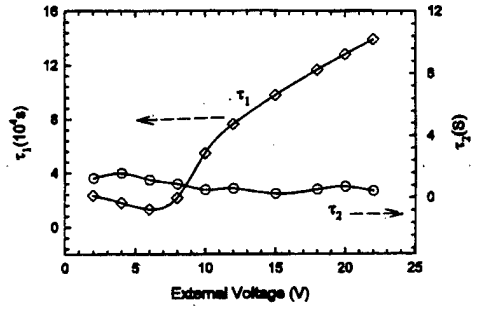


Fig. 10. The fitted lifetime τ_1 (♫) of excited protons and the rate constant τ_2 (♦) according to Eq. 11 as functions of the external voltage. The flash power and flash period are 129.0 mW and 10 s, respectively.

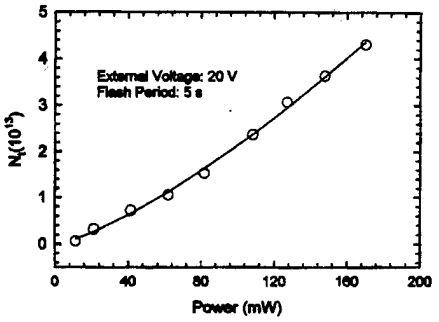


Fig. 11. The number of protons generated at the time when the flash laser is turned off, determined by integrating the fitted photocurrent curves with respect to time, as a function of flash power. Symbols indicate the integrated result and the solid line shows the fitted result. The external voltage is 20 V and flash period is 5 s.

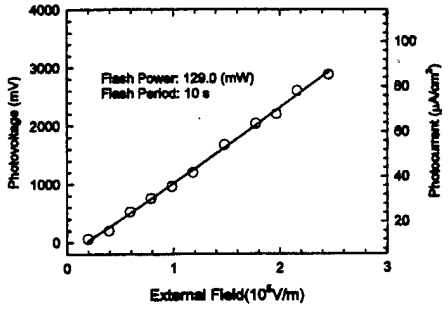


Fig. 12. The fitted average displacement of excited protons as a function of flash power according to Eq. 11. One can see that it is only slightly dependent on power. This may indicate that the average displacement is mainly determined by the composition of the sample.

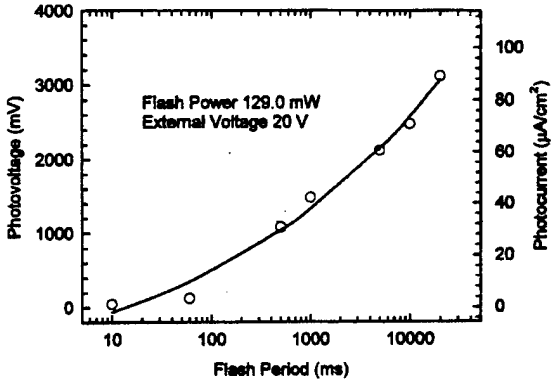


Fig. 13. Photovoltage and photocurrent as a function of external field. Symbols indicate the measured result and the solid-line shows the fitted result.

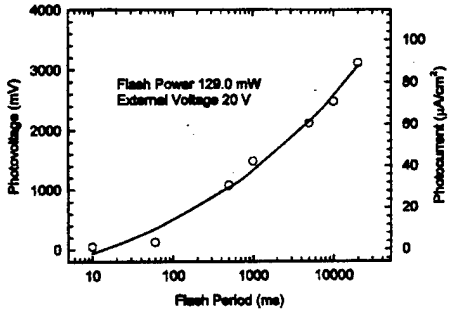


Fig. 14. Photovoltage and photocurrent as a function of flash period. Symbols indicate the measured result and the solid line shows the fitted result.

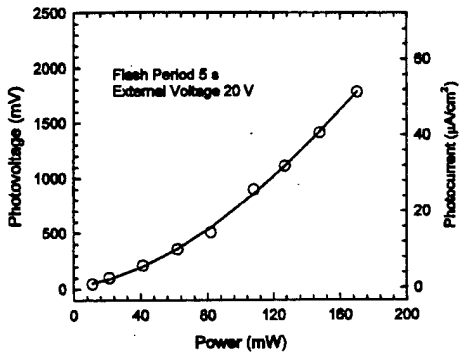


Fig. 15. Photovoltage and photocurrent as a function of flash power. Symbols indicate the measured result and the solid line shows the fitted result.

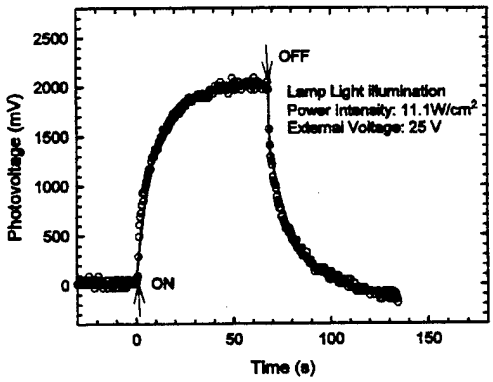


Fig. 16 Photovoltage as a function of time with CW lamp light illumination. External voltage is 25 V and illumination intensity is 11.1 W/cm^2. Symbols indicate the measured result and the solid-line shows the fitted result. One can see that the sample also gives strong response to illumination by incoherent light.

4. Importance

This work is important to the formation of a multifunctional display material. It has broader implications for biomimetic materials synthesis and optical detector materials and devices.

5. Activities During Quarter

We participated in the related activities during the prior program quarter.

Journal Papers

1. L. Zhang, T. Zeng, R. Claus, K. Cooper, Y. Liu and F. Zhang, "Electro-optic property measure-ments of electrostatically self-assembled ultrathin films," accepted for publication in Optics Communications, November 2000.

2. T. Zeng, R. Claus, Y. Liu, F. Zhang, W. Du and K.L. Cooper, "Piezoelectric ultrathin polymer films synthesized by electrostatic self-assembly processing," accepted for publication in Smart Mater. Struct., December 2000.

Conference Papers

1. W. Zhao, R.O. Claus, K. Cooper, Y.Liu, F.J. Arregui, "Self-Assembled Nanostructured Optical Fiber Temperature Sensors," OFS-14, Venice, October 2000.

2. F. Arregui, I Matias, Y. Liu, K. Cooper and R. Claus, "A Dichlormethane Gas Optical Fiber Sensor Fabricated Using ESA Monolayer Method," OFS-14, Venice, October 2000.

3. F.J. Arregui, I.R. Matias, K.L. Cooper, R.O. Claus and Y. Liu, "Optical Fiber-based Sensor of Harmful Gas Fabricated Using the Electronic Self-Assembly Monolayer Process," submitted to Optical Fibers and Sensors for Medical Applications, BIOS, San Jose, 2001.

Key Awards

1. Richard Claus received recognition as Virginia's "Scientist of the Year" for calendar year 2000.
2. Richard Claus has been nominated for membership in the National Academy of Engineering.

Visits During Reporting Period

1. Richard Claus visited WPAFB on 28 November 2000 to discuss display technology progress
2. Richard Claus visited Lockheed Martrin Aeronautics in Marietta Georgia on 10 January 2001 to discuss the program and cooperation.

6. Plans for Next Quarter

During the next quarter we plan to return emphasis to our local area network activity and plan to report on progress underway in that work.

7. Issues

No issues have arisen during the quarter.

Task 2.5

Crane Ship Simulator – Status Report

Abstract

A crane ship simulator in a Computer Automated Virtual Environment (CAVE™) is being developed at Virginia Tech. A motion platform will be used to provide motion queuing to the user of this simulator. The user will be immersed in a virtual environment, with a four-wall CAVE™ (which has three vertical walls and a floor), motion base with operator seat, control console with joystick crane controls, and four-channel sound. (http://thor.sv.vt.edu/crane/)

General virtual prototyping system

The crane ship simulator system is being constructed with an existing CAVE™ system. Since this CAVE™ is being used by many other unrelated projects, the development of the crane ship simulator is driving the development of a general virtual prototyping system due to necessity, in order for this crane ship simulator to co-exist with other unrelated projects in this CAVE™ and motion based system. The crane operator seat and console assembly is a single piece and has a mass of less than one hundred kilograms, so that it can easily be removed or replaced with a different unit in a matter of minutes. All other hardware components of the crane ship simulator are not specifically made for just the crane ship simulator and can be used with other projects. The crane operator seat and console assembly may also be used with other projects. This CAVE™ and motion based system has the unique ability to be reconfigured in minutes. The modular structure of the hardware and software of this system enable it to be a general virtual prototyping system as any CAVE-like system, but with the additional feature of user motion queuing. It is believed that the motion will make the CAVE™ an even more immersive environment, which will make it a more effective tool.

Virginia Tech CAVE™

Significant progress has been made on constructing the CAVE™ floor support structure for the motion platform and CAVE ™.

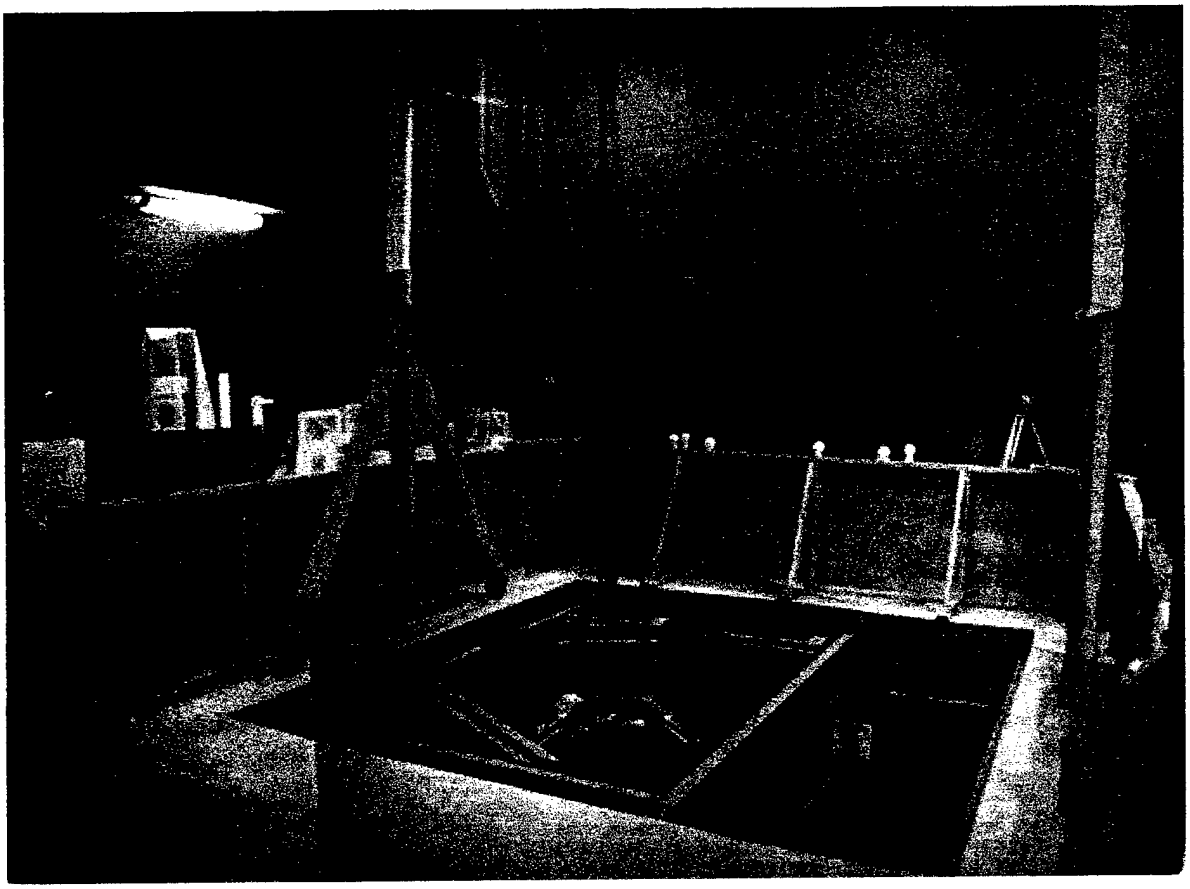


Figure 1 January 16, 2001, Room 3090 Torgersen Hall, Virginia Tech: a rough composite image of the MOOG motion base being lowered into the floor.

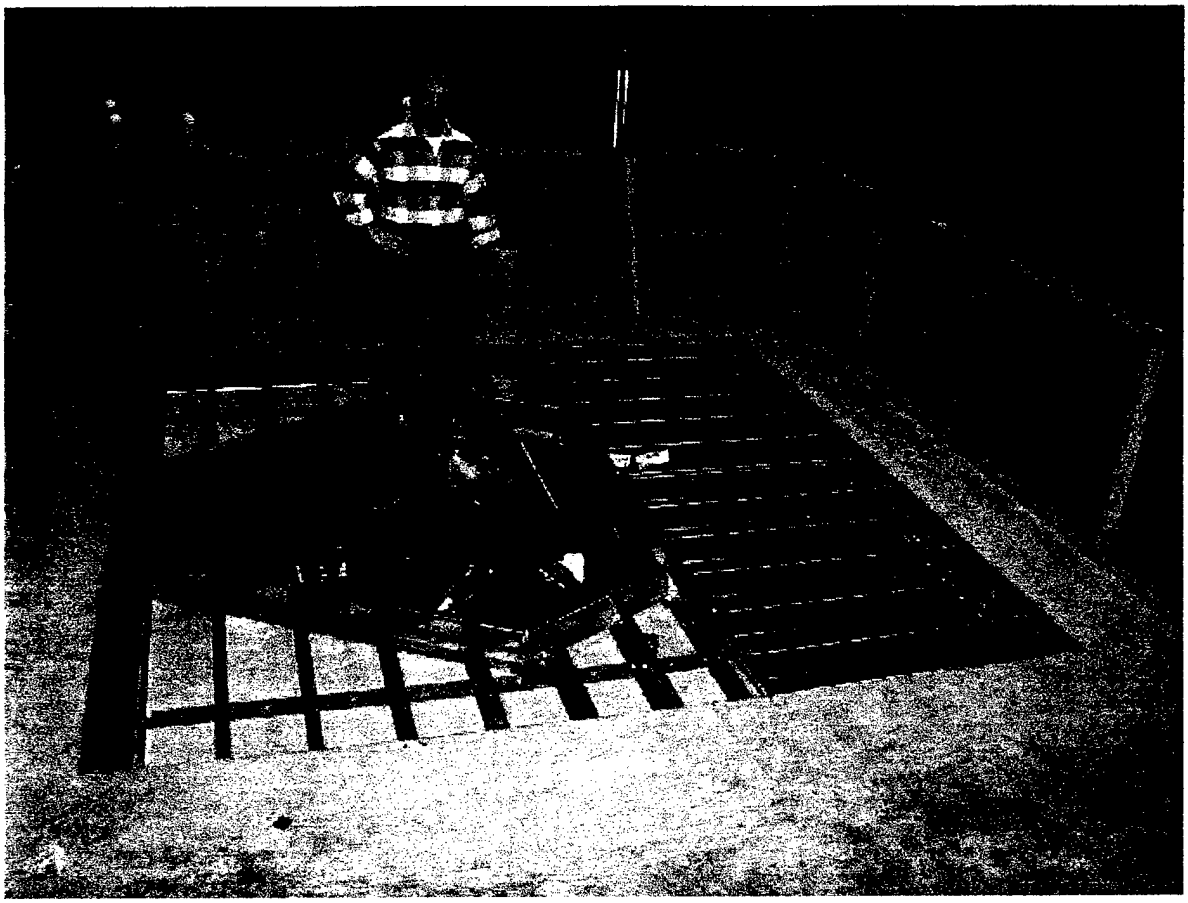


Figure 2 February 6, 2001, Room 3090 Torgersen Hall, Virginia Tech: the floor with joists at the time of the writing of this report.

The construction of the building floor is completed. We project that construction of the CAVE™ will be completed around the first week of March 2001.

Software Design

The simulator system structure is a cluster of cooperating programs. These programs are coupled to each other by fixed data structures that retain the current state information of the system in interprocess shared memory from which the programs can read and write. This design facilitates a modular structure for the system. All the programs in the simulator run asynchronously with respect to all other programs. Some programs consist of more than one process or thread. Some programs are simulating physical dynamical models whose cyclic solver is synchronized to real-time.

The DIVERSE Toolkit (http://www.diverse.vt.edu/DTK/)

The DIVERSE Toolkit (DTK) is the "glue" that ties together the programs in the simulator. Briefly, DTK consists of the server program, a C++ client application programming interface (API), and small utility programs. The DTK server manages the interprocess shared memory, provides an interface for other programs to serial hardware devices via shared memory, and a seamless interface to IP networked shared memory. All programs in the simulator use the DTK C++ client API. The DTK server is central to the system in that it is the only program required to be running at all times for the simulator to be running. All of the other programs (modules) can be developed and run independently of most other modules. Modules can be emulated and started and stopped without corrupting other running modules. When the system is running in a steady state the primary interprocess communication (IPC) mechanism used is DTK interprocess shared memory.

The largest difference between this programming paradigm and one using message passing interface (MPI) is that the programs distributed in this system typically run asynchronously. The network IPC methods in DTK are usually unreliable (i.e., UDP/IP). However, reliable TCP/IP is used for changing the system configuration, like adding an observer to the system or communicating discrete events, such as turning on a light. The UDP/IP methods are used for transferring most of the network information, such as when a model of a hydraulic cylinder is continuously expanding and contracting due to an operator's input. The changes in the cylinder's length are being fed to the network continuously at a regular rate. If a cylinder length network IP packet is lost, there will not be a need to send again that information because, in most cases, the next cylinder length packet will come before a replacement IP packet can arrive. This method is analogous to the way real systems interact. For example, if you blink your eyes, the light that your eyes did not receive is not sent again, but instead the latest light information is sent to your eyes when you reopen them.

Integration design of programs (modules) in the simulator system

The simulator programs contain DTK client objects that enable the making of DTK shared memory objects, which in turn enable communication between the programs in the system. The code in the programs may be written in any combination of languages that may be linked with C++ objects, which includes FORTRAN, C, and C++.

Most of the programs in the system, after being initialized, run in a cyclic steady-state simulation loop. For this case the programs code may be split into two logical functions (subroutines or methods), "initialization"

and "cyclic advance." In the case of a C++ implementation, a class constructor may be used for initialization. A possible third function can be "closure" for the removal of shared computational resources that may have been allocated when the program runs, so that the program may call this function before exiting. We refer to the code of these functions as "task functions." The writing of the task functions does not require the use of the DTK API. For programs that must run in real-time, the real-time synchronizing code is not included in these task functions, but is introduced outside of these task functions, so that the code of these task functions may be used on other platforms and in other applications. and testing the code is easier. The task functions provide enough output variables so that any needed changing system-state information, such as position, angular acceleration, and time, can be gotten from the output variables. The task functions provide enough input variables so that all changing inputs are passed to the task functions. The code of these task functions may preserve the dynamical state of its task or the dynamical state of its task may be passed to the task functions as a combination of input/output variables. The latter case may be more resource (CPU and memory usage) efficient in some cases. When possible, the cyclic advance task function should allow the user to specify the time to advance to.

List of Simulator System Modules (running programs)

All of the simulator modules will continually evolve. The modules marked as done are currently in a usable state.

1. DTK: server system service provider (done)
2. Render: drawing visual (done)
3. Wave: compute wave height and ship's x and y position relative to the ocean grid (done)
4. LAMP: (http://www.ship.saic.com/) ship hull dynamics model (done)
5. Fun Filter: visual and motion base relative motion filter
6. Crane Model: crane dynamics model
7. Crane Controller filter operator input to crane model
8. Motion Base Controller is a controller for engaging and shutting down the MOOG motion base (done)
9. Sound simulation activated sound

NAVCIITI Task 3.1: Network Protocol Interoperability
Report (November 2000-January 2001) Status Report

February 22, 2001

3.1. Network Protocol Interoperability

3.1.1. Background

Task 3.1, Network Protocol Interoperability, focuses on enabling interoperability between heterogeneous networks that may belong to and be managed by different organizations, including allies and coalition partners. Primary technical areas being investigated as part of this task include distributed management of a multi-organizational virtual network; implementation and dynamic management of sets of policies enabling differentiated quality of service (QoS); routing in mobile and ad-hoc networks; wireless network security. We are currently building a distributed network test bed, with a presence on the Virginia Tech main campus in Blacksburg and at the Alexandria Research Institute. This test bed will support experimental investigations into each of the four areas mentioned above. During Year 2 (the first year of this particular task), the focus is on deploying the test bed and developing and evaluating specific interoperability solutions.

3.1.2. Planned Effort

Efforts planned for November 2000-January 2001 were to (i) deploy portions of the testbed, (ii) refine and implement the distributed network management model, (iii) develop techniques for emulating a dynamically changing wireless network topology, and (iv) produce an initial high-level model of the QoS policy distribution and enforcement. We had planned to present results-to-date from these efforts, especially item (ii) above, to our Navy contacts.

3.1.3. Accomplishments

As noted above, principal accomplishments during the period November 2000-January 2001 were to (i) deploy portions of the test bed at networking labs in Blacksburg and at the Alexandria Research Institute, (ii) refine and implement the distributed network management model, (iii) develop techniques for emulating a dynamically changing wireless network topology, and (iv) develop an initial model of QoS policy distribution and enforcement.

The principal investigator, Scott Midkiff, and one of the graduate research assistants in the project, Erik Hia, participated in the Virtual Operations Network (VON) Applications and Security Workshop held in November 2000 with the objective of discussing joint

US/UK VON program goals for 2001 and beyond. During this meeting, we briefed our Navy contacts and UK partners on our proposed architecture for distributed management of heterogeneous networks and how this architecture can be incorporated into experiments and demonstrations currently planned under the VON effort.

We present brief summaries of the progress being made in each technical area.

- Mobile network routing. We continue our investigation of traditional OSPF and of emerging mobile ad hoc network (MANET) routing protocols for routing between mobile platforms (ships) in the VON backbone. Our experiments will account for the fact that mobile network topologies are dynamically changing, with links becoming unavailable and new links being established. We have developed a technique for controlling the links in the test bed topology through the use of firewalls; traffic management functions (such as shaping) and the emulation of link errors are currently being added.

- Quality of service. We have developed an initial architecture for the distribution and enforcement of policies to support QoS. This architecture makes use of Local Policy Decision Points (LPDPs) that are co-located with Policy Enforcement Points (PEPs) in order to more efficiently handle service requests and validate application-level traffic classification. We will demonstrate the use of the policy architecture in support of differentiation mechanisms capable of providing preferential treatment to high-priority traffic. These mechanisms include IP DiffServ (Differentiated Services) and IP DiffServ with RSVP (Resource Reservation Protocol).

- Network security. We are examining the network overhead of using IPSec and will demonstrate quantitative results. A taxonomy of security attacks has been developed, with particular focus on attacks on wireless networks.

- Network management. We are investigating and will demonstrate the use of management schemes based on SNMP that attempt to reduce backbone capacity consumed by management traffic. These management schemes take a hierarchical approach, with the use of midlevel and high-level management agents.

We continued to acquire equipment as part of the planned test bed that will support experimental studies of layer 3 (network) and higher layer protocols, with some capabilities for examining layer 2 (data link) protocol issues. Some of the test bed is currently deployed and in use for initial investigation of routing and management issues. We have started occupying the space at the Alexandria Research Institute that will house the portion of the test bed located in Northern Virginia; first equipment for that location was acquired, with more equipment currently being ordered.

We continue to work with Task 3.2 personnel to identify and pursue areas of potential collaboration.

3.1.4. Importance

The activities in Task 3.1 relate to the general problems encountered in a Virtual Operations Network. The VON will likely rely on a backbone built from low data rate wireless links and having a dynamic topology. The investigation of the performance of the OSPF routing protocol and MANET routing protocols will provide metrics critical to the selection of an appropriate routing strategy, protocol, and configuration for the VON and for understanding overhead for routing. Bandwidth management and mechanisms for differentiated services across the backbone will be critical to ensuring that important application data is transferred in a timely manner. Our exploration of QoS will lead to recommendations to achieving these capabilities. Network management will be critical to the rapid deployment and operational maintenance of shared network resources. Numerical results produced during the past quarter can guide selection of a security strategy for network management, while new work is investigating a management architecture to reduce traffic on low data rate links. Security will be critical to all aspects of the VON's network infrastructure. Our investigation will characterize alternative security approaches and will provide an understanding of network overhead introduced by these approaches.

3.1.5. Activities

Virginia Tech faculty investigator and student participated in the VON Applications and Security Workshop in Arlington on November 30, 2000.

3.1.6. Plans for the Next Quarter

We will finish testing the set up of full subnet of the test bed and order remaining equipment to make other subnets operational. We will continue to investigate demonstration areas, implement experimental systems, and develop methods to validate data for Task 3.2 explore ways that scenarios from Task 3.2 can be used to drive test bed experiments. We will participate in briefings with our Navy contacts and their industrial partners to identify work synergies. We will provide the Navy with a technical report with detailed results and discussion of the technical findings to date.

3.1.7. Issues

There are no known issues at this time.

3.2 Network System Interoperability

Background. This task is in direct support of the Virtual Operations Network (VON) Project, and the research team is working closely with Navy scientists and engineers at NSWCDD and SSC-SD The project receives overall direction from LCDR Dave Jakubek of ONR. The VON project has the goal of establishing interoperability among a coalition task force comprised of platforms with varying communications and computing capabilities. The Virginia Tech Systems Research Center (VTSRC) team is responsible for developing the models and algorithms to establish mission feasibility of performing the assigned tasks given the communications assets provided by the participants. A future goal is to provide decision support to the coalition command that will enable dynamic evaluation of communications capability in response to changes in mission or coalition assets.

Planned Effort for the Quarter. The focus of effort for this period has been on the development of data requirements for the communications feasibility determination that is at the core of the static model. In meetings and conference calls, the communication needs of a VON coalition force has been clarified, with the primary objective of assessing the level of timing granularity required to produce a valid determination by the algorithm. This effort has forced the VTSRC team to define the role of Coalition Mission Planner (CMP). This role has the primary responsibility for specifying the Navy Tactical Actions (NTAs) that are submitted to the Naval Architectural Database (NAD). The output from the NAD is in the form of Information Exchange Requirements (IERs). The CMP must examine the IERs in comparison with coalition mission objectives to effect a transformation into messages with a temporal ordering (a chronology). The result must define message exchanges at the proper level of detail to support a valid determination of adequate communications resources to realize the mission objectives.

Status of Effort. Based on feedback from meetings in the fall, the definition of a tentative system architecture has been accomplished. The three objectives stated for this quarter have been realized, but the long holiday period during December has not permitted further interchange of ideas with the VON team. We now have (1) the data requirements for the algorithm, (2) the points constituting a draft definition of the mission planning function, and (3) a draft representation of the primary algorithm to determine feasibility. We also have established the data dictionary and begun the development of requirements for decision support software needed by the Coalition Mission Planner (CMP) in accomplishing that function.

Following our strategy conveyed in the last quarterly report, we are sharing plans and progress reports with the Tasks 3.1 research team headed by Professor Midkiff.

Importance of Objectives. The VON Project is in direct response to a report of the Navy Research and Advisory Council (NRAC) calling for interoperability in communications among participating nations in a coalition operation. The VTSRC support entails analytical models for feasibility assessment during mission planning. This effort requires the development of a static model to assist in the mission planning function. A dynamic model providing decision support during mission execution is envisioned as future work. A briefing of the modeling role and the development of models and algorithms is anticipated in preparation and planning for a major demonstration in the summer 2001 timeframe.

Activities. A briefing of the communications database architecture, the distinction in roles of the static and dynamic models, and an overview of the algorithm for feasibility assessment within the mission planning function was delivered to the joint US and UK team on November 28, 2000. At that time the relationship between Tasks 3.1 and 3.2 was described.

Plans for the Next Quarter. The objectives for the next quarter are:

(1) to layout specific data record formats for the Coalition Operations Database,
(2) to expand and refine the description of the mission planning function, giving special attention to the requirements for supportive software,

(3) to produce a high-level specification of the static model, including a detailed specification of the algorithm for feasibility assessment, and
(4) to design an implementation (program) of the algorithm for feasibility assessment.

Issues. The VTSRC research team concerns continue regarding our ability to meet the VON demonstration schedule given funding cuts imposed at the NAVCIITI project level. The increase in funds for Year 3, to offset the reductions in Year 2 to account for interim funding allocations, is desperately needed to increase the effort to the level specified in the original task statements.

Task 4.1

1. Introduction

This quarterly report of NAVCIITI Year Two describes the objectives, proposed research, current status, related activities, and planned efforts of Task 4.1 "Real-Time Resource Management" that was performed during the period November 1, 2000 through January 31, 2001. The organization of the report is as follows: We describe the background and objectives of our proposed research in Section 2. A description of the proposed research and our planned efforts is given in Section 3. We summarize the current status of the work in Section 4. Section 5 describes a new resource management algorithm that we have developed during the period November 1, 2000 through January 31, 2001. We discuss the publication, presentation, and other related activities during this quarter in Section 6. Finally, we conclude the report by describing the planned efforts for the next quarter in Section 7.

2. Background and Objectives of Proposed Research

Real-time computer systems that are emerging for the purpose of strategic mission management such as coordination of multiple entities that are manufacturing a vehicle, repairing a damaged reactor, or conducting combat are subject to great uncertainties at the mission and system levels. The computations in the system are "asynchronous" in the sense that processing and communication latencies do not have known upper bounds and event arrivals have non-deterministic distributions. Computational parameters such as task execution times, communication delays, periods, and event arrivals are non-deterministic as they are dependent upon the conditions of the external environment where the system is deployed and the operational scenario of the system that is situation-specific.

Such real-time mission management applications require decentralization because of the physical distribution of application resources and for achieving survivability in the sense of continued availability of application functionality that is situation-specific. Because of their physical dispersal, most real-time distributed computing systems are "loosely" coupled using communication paradigms that employ links, buses, rings, etc., resulting in additional uncertainties e.g., variable communication latencies, regardless of the bandwidth. The characteristics of the application combined with the physical laws involved in distribution thus contribute to the non-deterministic and asynchronous behavior of the system [Jen92, Jen99].

New advances in real-time distributed systems research [Quo97] have produced quality-of-service (QoS) technologies that allow asynchronous real-time distributed systems to specify and negotiate service expectations such as timeliness and survivability, which was not previously possible under classical hard real-time computing theory. However, the QoS management literature or the classical real-time literature does not address three fundamental problems in asynchronous real-time distributed systems.

First, the literature does not provide solutions to the problem of achieving the real-time requirements of *aperiodic* computations in asynchronous real-time distributed systems. Hard real-time computing theory provide scheduling and resource management solutions for aperiodic computations primarily in static, centralized systems where the computational parameters such as execution times, communication delays, and (distributions on) event arrivals are known with absolute certainty. On the other hand, the QoS management technologies focus primarily on periodic computations, where adaptation of the periodics to the "right level" of their desired QoS is dynamically performed to accommodate run-time uncertainties and workload changes. It is difficult, if not impossible to apply the QoS techniques for aperiodics. This is because it is impractical to model the real-time requirements of aperiodic computations as multiple levels of service that vary in their quality of computational results and requirements of resources. Furthermore, allocation of resources (by adaptation or negotiation) for aperiodic computations *at event arrival* is practically infeasible due to their high cost.

Second, the classical real-time communication literature focuses primarily on static systems, where the communication parameters of the application such as message latencies, message periods, and message arrivals are deterministically known. Further, the QoS literature is sparse with works on dynamic adaptation mechanisms at the trans-node *message-level*. Rather, the QoS technologies primarily focus on process-level adaptation mechanisms such as process replication and imprecise computations. We believe that dynamic adaptation mechanisms at the message-level can produce significant benefits. Furthermore, such adaptive communication techniques have great practical significance with the emerging communication standards such as IEEE 802.1p [IEEE802a, IEEE802b] that supports message prioritization on Ethernet networks. This facilitates a mechanism for performing run-time adaptation (at the message-level) on such networks that are cheap and widely available.

Third, the performance of the real-time QoS management algorithms that appear in the literature is only *empirically* studied for the most part. The performance of the QoS mechanisms is sometimes studied through simulation or they are implemented as part of system software layers (e.g., middleware) and performance characterization is done using

real-time benchmarks such as [MV+96, SWR99]. This certainly validates the effectiveness of the techniques. However, the *general feasibility conditions* of the QoS techniques for satisfying the real-time requirements are unknown. The fundamental limitations of the QoS management techniques and the boundaries of satisfiability of the requirements are not known. This is a difficult, but fundamental problem since there does not exist a theory for designing QoS management algorithms and for reasoning, evaluating, and validating their behavior under all possible operating conditions of the system.

We propose fundamental technology that solves each of the three problems outlined above. The proposed technologies include (1) *proactive resource management algorithms* that satisfy the real-time requirements of aperiodic computations, (2) *adaptive communication techniques* that adapt the application to workload fluctuations through trans-node message-level mechanisms and satisfy their real-time requirements, and (3) *hybrid feedback control theoretical techniques* that allow real-time system engineers to mathematically design adaptive resource management techniques and reason about their behavior under all possible operating conditions of the system.

3. Description of Proposed Research and Planned Efforts

In this section, we discuss our proposed solution approach for proactive resource management of aperiodic computations. Our proposed solution approaches for adaptive communication and hybrid feedback control theoretical techniques were presented in the quarterly progress report (number 13) for the period August 1, 2000 through October 31, 2000. This quarterly report (number 13) also reported:

(1) the state-of-the-art of each of the research problems that we are addressing—proactive resource management, adaptive communication, and hybrid feedback control theory—to illustrate the significance of the problems

(2) adaptive communication mechanisms that we had developed for IEEE 802.5 (Token Ring) and 802.1p (Ethernet) networks that support priority-based message scheduling, and FDDI networks that use the Timed Token protocol, their experimental evaluation, and performance results

(3) a hybrid modeling framework for developing control-theoretical solutions for adaptive resource management problems.[1]

Section 3.1 reviews our proposed solution approach for proactive resource management of aperiodic computations.

[1] Hybrid feedback control-theory-based resource management techniques will be developed during Year 3.

3.1 Proactive Resource Management of Aperiodic Computations

We propose proactive resource management algorithms for aperiodic computations that forecast their resource needs, given the current internal resource availability and desired external load. We propose to develop forecasting techniques that predict the end-to-end resource requirements of aperiodic computations. Furthermore, we propose to model the desired number of aperiodic events that needs to be satisfied by the system in a timely manner as the desired QoS, as opposed to the accuracy or quality of the computations. Thus, the application will desire a certain number of aperiodic events that it would prefer to be "ready for" at all times. The proposed proactive adaptation mechanism will forecast the resource requirements that are needed to satisfy the desired number of aperiodic events, given the current resource availability, workloads, and real-time requirements of other, possibly, periodic tasks.

The forecasted requirements will be used to drive resource allocation strategies. For example, if the forecasted resource requirements are not currently available, then resource allocation is performed through adaptation mechanisms such as replication of processes of the aperiodic computation. Resource allocation strategies will determine (1) candidate processes of the aperiodic computation that needs to be replicated, (2) number of replicas that are needed that can satisfy the desired real-time requirements, and (3) processors that are required for the computational needs of the replicas that will satisfy the desired requirements.

The proactive resource allocation algorithms will determine whether it is possible to satisfy the real-time requirements of the aperiodics through dynamic adaptation and without jeopardizing the timeliness of other computations, given the current situation. Furthermore, if a feasible resource allocation is possible, the algorithms will determine the resource allocation. If the algorithms determine that it is not feasible to achieve the collective requirements of the application, then negotiation strategies will be employed to achieve the optimal satisfaction of real-time requirements across computations. Thus, proactive resource management algorithms perform resource allocation and negotiation based on forecasted application performance to achieve the desired requirements.

4. Current Status and Accomplishments

We have developed several proactive resource management algorithms for aperiodic tasks that achieve their real-time requirements. Two of the resource management algorithms use the well known slack stealing and round-robin processor scheduling algorithms, respectively, for scheduling application programs on processors. We had developed these resource management algorithms during the period August 1, 2000

through October 31, 2000. The algorithms, their experimental evaluation, and performance results were reported in our quarterly report (number 13) for the work performed during that period. So they are not discussed here.

As discussed previously, we have also developed a dynamic message reprioritization algorithm for IEEE 802.5 (token ring) and 802.1p (Ethernet) networks that support the priority-driven protocol and an adaptive communication algorithm for the Timed Token protocol. The adaptive communication algorithms were developed during the period August 1, 2000 through October 31, 2000 and were reported in our quarterly report (number 13) for the work performed during that period. So they are also not discussed here.

We have developed a new proactive resource management algorithm for aperiodic computations that use *best effort* scheduling algorithms for scheduling application programs on processors. We discuss the new algorithm in Section 5. Similar to the resource management algorithms that we have developed earlier, we study the performance of the algorithm through a combination of benchmarking and simulation.

5. A Best-Effort Resource Management Algorithm for Aperiodic Computations

We present a best-effort resource management algorithm for satisfying the real-time requirements of aperiodic tasks in asynchronous real-time distributed systems. The resource management algorithm uses the DASA [Cla90] and the LBESA [Loc86] best-effort scheduling algorithms for scheduling application subtasks on processors and for routing application messages at routers. The resource management algorithm is "best-effort" in the sense that it makes a best effort—as specified by the application—to adapt the application to workload uncertainties to maximize system-wide benefit and system-wide timeliness.

We use Jensen's benefit functions [Jen93] for describing application timeliness requirements. Thus, the urgency of each computation of the application is expressed in terms of the benefit it provides to the system as a function of the time at which the computation is completed. Furthermore, we extend Jensen's benefit accrual model with the notion of *adaptation functions* for expressing the anticipated workload scenarios of the application. The adaptation functions are used to describe the anticipated application workload as a function of the time at which it is anticipated to occur. The origin of the axes of the function is the current time and is treated by the resource management algorithm as a "resource management event" in the sense of the start time of an expected scenario for the workload. This is analogous to a scheduling event. The function is

specified for a fixed duration of time into the future and ends at a time instant called the "time horizon."

For example, the anticipated rate of event arrivals that trigger aperiodic computations and the anticipated workload of periodic computations per task period are specified as a function of their occurrence times, relative to a start time. Two example adaptation functions are shown in Figures 1(a) and 1(b).

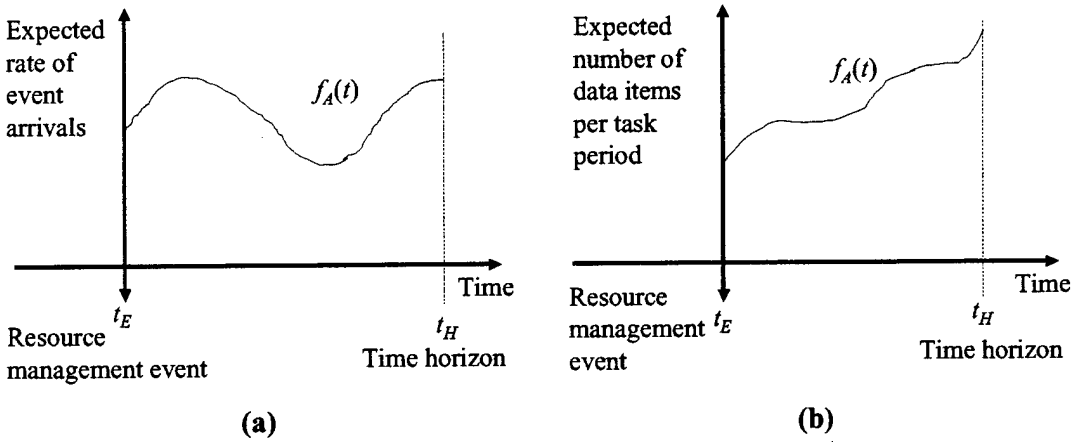


Figure 1. Adaptation Functions for (a) Aperiodic Tasks and (b) Periodic Tasks

The adaptation functions are application-specific, defined system-wide, and are dynamically modified by the user as the anticipated workload scenarios change. When changes in the anticipated (periodic or aperiodic) workloads are specified and triggered by the user, the proposed best-effort resource management algorithm forecasts the resource needs that will maximize the aggregate system benefit and aggregate application timeliness during the anticipated workload situation. The forecasted resource needs are then allocated to the application. We use the mechanism of replication of subtasks of end-to-end tasks as the adaptation strategy. Thus, the proposed best-effort resource management algorithm seek to provide the best aggregate benefit and best aggregate timeliness to the application that is possible under the current application and resource situation.

Section 5.1 discusses the application and system model that is used by the algorithm. We then present the algorithm in Sections 5.2 through Section 5.6. The experimental evaluation of the algorithm is discussed in Section 5.7.

5.1 The Application and System Model

The application and system model that we have assumed in this work is based on the Anti-Air Warfare (AAW) system of the U. S. Navy [WRSB98]. We introduce some simple notations to describe the application characteristics. The notations are later used in the presentation of the algorithm. The application properties and the notations are described as follows:

1. The set of tasks in the application is denoted by the set $T = \{T_1, T_2, T_3 ...\}$.
2. A task T_i can be either periodic or aperiodic. Each aperiodic task T_j has a "triggering" periodic task T_k that triggers its execution. After T_k completes its execution, it generates a set of events that trigger the execution of T_j.
3. The period of task T_i is denoted as *period*(T_i), if T_i is periodic. If T_i is aperiodic, *period*(T_i) denotes the period of the periodic task that triggers task T_i.
4. Each task T_i is assumed to consist of a set of subtasks (executable programs), which execute in a *serial* fashion. We use the notation $T_i = [st_1^i, m_1^i, st_2^i, m_2^i, ..., st_n^i, m_n^i]$ to represent a task T_i that consists of n subtasks and n messages to be executed in series. That is, $st_i^k (i > 1)$ cannot execute before message m_{i-1}^k arrives.
5. The set of subtasks of a task T_i is denoted as $ST(T_i) = \{st_1^i, st_2^i, st_3^i, ..., st_n^i\}$ and the set of inter-subtask messages of the task T_i is denoted as $MS(T_i) = \{m_1^i, m_2^i, m_3^i, m_4^i, ..., m_n^i\}$.
6. *type*(st_j^i) is a function defined over the subtasks that identifies whether the parent task T_i of a subtask st_j^i is periodic or aperiodic. The function returns one of the values, *PERIODIC* or *APERIODIC*.
7. An important assumption that is made regarding subtasks is that they can be *replicated* at run-time. The idea behind replication of subtasks is that once a subtask is replicated, the replicas of the subtask can share the workload that was processed by the original subtask. Further, concurrency can be exploited by the executing the replicas on different processors and thereby the end-to-end latency of the task can be reduced. Thus, replication is allowed as a means to reduce task latencies and improve task timeliness when task workloads increase at run-time.
8. The states of the replicas of the subtasks and their consistency is not addressed in this work, as we assume that the tasks process data objects that are "continuous" in the sense that their values are obtained directly from a sensor in the application environment, or computed from values of other such objects. The replicas are thus assumed to be *temporally consistent* (e.g., sufficiently up-to-date) without applying every change in value, due to the continuity of physical phenomena.

9. The estimated execution time of the subtask st_j^i for processing data of size d on a processor with utilization u is denoted by $eex(st_j^i, d)$.
10. The estimated communication delay of the inter-subtask message m_j^i that carries data of size d is denoted by $ecd(m_j^i, d)$.
11. The deadline of task T_i is denoted by $dl(T_i)$.
12. The application hardware is assumed to consist of a set of distributed processors that are interconnected through a router in a star-topology fashion. There is a full-duplex Ethernet link (IEEEE 802.3) between each node and the router. The hardware is assumed to consist of a set of m processors denoted by the set $PR = \{p_1, p_2, p_3, ..., p_m\}$. The processors are assumed to be homogenous and each processor is assumed to have a piece of private memory that can only be accessed by the processor itself. Further, the clocks of the processors are synchronized using an algorithm such as [Mills95].
13. The scheduling algorithm that is used at the router to schedule messages that compete to get to the same destination node is the same as the scheduling algorithm that is used to schedule subtasks on processors.
14. The processors that are executing a subtask st_j^i at time t is denoted by $pr(st_j^i)$.
15. The processor that is executing a subtask replica $st_{j,k}^i$ at time t is denoted by $prr(st_{j,k}^i)$.
16. The benefit of a task T_i is denoted by $V(T_i)$. We assume rectangular benefit functions for all tasks as in [Cla90]. Thus, completing a task anytime before its deadline will result in uniform benefit; completing it after the deadline will result in zero benefit.
17. The anticipated workload of a task T_i during a period p is given by $Adapt(T_i, p)$. The anticipated workload is defined in terms of the number of data items for periodic tasks and in terms of number of events for aperiodic tasks.

5.2 Overview of the Best Effort Resource Management Algorithm

As mentioned previously, the adaptation functions for a future window of time are given as input to the resource management algorithm. The functions are defined for each application task and they simply describe the anticipated workload for each task period for the future time window. The functional objective of the resource management algorithm is to determine (1) the number of replicas of the subtasks of each task for each task period in the future time window and (2) processors for executing the replicas. The performance objective of the algorithm is to (1) maximize the overall value accrued—measured as the sum of the value accrued for each of the tasks—during the future time

window, and (2) to minimize the number of (end-to-end) deadlines missed during the future time window. The algorithm determines the number of replicas and processors for executing the replicas for each subtask (of every task) for every period in the future time window.

The algorithm makes an initial processor assignment for the subtasks with the assumption that a single replica for each subtask is enough to satisfy the task timeliness requirement. When a resource management decision is triggered at run-time as a result of the user input (with adaptation functions), the algorithm first sorts all tasks according to their values. The algorithm then determines the number of replicas for each subtask and processors for executing the replicas that will satisfy the task deadline, starting from the highest value task. Note that we are assuming rectangular benefit functions for the tasks. Thus, the algorithm automatically maximizes the value accrued for the task, when it determines the number of replicas that will satisfy the task deadline, since the value accrued is zero if the task were to complete after the deadline.

Whenever a task T_i is considered for replication, the algorithm ensures that the timeliness of any tasks that have higher values than that of T_i are not affected. In other words, if the algorithm determines that it is not possible to satisfy the deadline of task T_i without starting enough replicas that will however, increase the completion times of subtasks of tasks having higher values than that of task T_i, then it stops replicating task T_i and considers the next (less value) task in the list. Thus, replication of a task is only performed up to a point where it will not affect the timeliness and thus the—aggregate benefit—of higher value tasks.

We now discuss the different components of the algorithm in the subsections that follow.

5.3 Deadline Assignment of Subtasks and Messages

To determine the number of replicas of each subtask, the algorithm breaks the end-to-end deadline of each task into individual deadlines for the subtasks and messages of the task. To assign deadlines to the subtasks and messages from the end-to-end deadline, we use a variant of the equal flexibility (EQF) strategy proposed in [KG97]. Observe that EQF requires knowledge of the execution times and communication delays to compute subtask deadlines. Therefore, the resource management algorithm uses estimates of the initial operating conditions of the system to derive the initial values of execution times and communication delays.

Let d_{init} denote the initial data size processed by a subtask st_i^k and transmitted by a message m_i^k, respectively. Let $eex(st_j^i, d)$ denote the estimated execution time of the subtask st_j^i for processing a data size d. Then, the deadline of the subtask st_i^k is given by:

$$dl(st_i^k) = eex(st_i^k, d_{init}) + \left(dl(T_k) - \sum_{j=i}^{m} eex(st_j^k, d_{init}) - \sum_{j=i+1}^{m} ecd(m_j^k, d_{init}) \right) \times \left[\frac{eex(st_i^k, d_{init})}{\left(\sum_{j=i}^{m} eex(st_j^k, d_{init}) + \sum_{j=i+1}^{m} ecd(m_j^k, d_{init}) \right)} \right] \quad (1)$$

The deadline of the message m_i^k is given by:

$$dl(m_i^k) = ecd(m_i^k, d_{init}) + \left(dl(T_k) - \sum_{j=i}^{m} ecd(m_j^k, d_{init}) - \sum_{j=i+1}^{m} eex(st_j^k, d_{init}) \right) \times \left[\frac{ecd(m_i^k, d_{init})}{\left(\sum_{j=i}^{m} ecd(m_j^k, d_{init}) + \sum_{j=i+1}^{m} eex(st_j^k, d_{init}) \right)} \right] \quad (2)$$

We describe the computation of the estimated execution time $eex(st_j^i, d)$ and estimated communication delay $ecd(m_j^i, d)$ in the following subsection.

5.4 Estimating Subtask Response Time

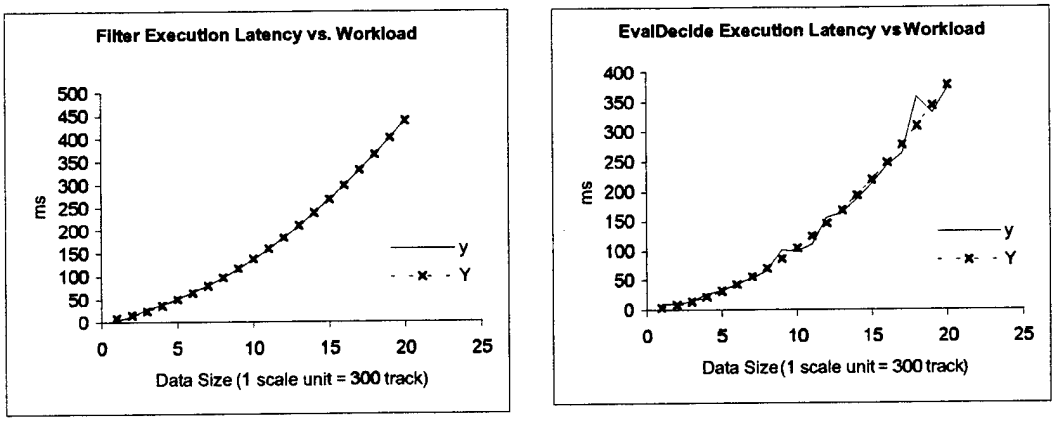
The response time of a subtask st_i^j of a task T_j under priority-based schedulers is given by the classical equation [ABR+93]:

$$R_i^j = C_i^j + I_i^j$$

where R_i^j is the response time of the subtask, C_i^j is the execution time, and I_i^j is the interference that the subtask experiences from other subtasks. When a subtask experiences such interference, the scheduler preempts it, and the subtask waits in the process ready queue until the scheduler selects it again.

We characterize the workload of a subtask as the number of data items that it needs to process, since the number of data items processed by the subtasks constitute the most significant part of the application workload in systems such as the Navy's AAW. Thus, we define the subtask execution time C_i^j as a function of the number of data items that it

needs to process. To determine this function that can provide a reasonable approximation of subtask execution latencies, we use the real-time benchmark [SWR99] as an example application. The benchmark is a functional approximation of the Navy's AAW system. We measure the execution time of the subtasks of the benchmark for a variety of data sizes without any interference from other application subtasks. The measurements are then used to define a regression equation that computes execution latency as a function of subtask data size.



(a) **(b)**

Figure 2. Execution Latencies of (a) Filter and (b) EvalDecide Subtasks Under Varying Data Sizes with no Contention

Figures 2(a) and 2(b) show sample plots of execution latency measurements of two application subtasks of the benchmark called *Filter* and *EvalDecide*. The measurements were made by dedicating the processor solely to the execution of the subtask under consideration. The solid lines (called "y" in the figure) show the real execution time and the dashed lines (called "Y" in the figure) show the second order curve that was computed for the real execution times. This second order curve is defined by the second order regression equation:

$$C_i^j(d) = a_i^j d^2 + b_i^j d \qquad (3)$$

where d is the data size in hundreds of data items, u is the CPU utilization in percentage, and a_i^j and b_i^j are constants that are dependent upon the application subtask.

To determine the response time of a subtask on a processor, we need to consider all subtasks that execute on the processor and belong to tasks that have a higher value than the parent task of the subtask under consideration. Furthermore, we need to determine the arrival times of the subtasks that share the processor with the subtask under

consideration. To determine the subtask arrival times, we make the following assumptions to simplify the analysis:[2]

(1) All periodic tasks arrive at the beginning of their periods
(2) The arrival time of an aperiodic task is considered to be the end of the period of the periodic task that triggers its execution. We make this assumption even though the actual triggering may happen earlier, when the triggering task finishes earlier than its deadline.

Based on the above assumptions, the arrival time of a subtask can be determined by adding the relative deadlines of all the subtasks and the messages that precede the subtask under consideration with the arrival time of the parent task of the subtask. For example, if the arrival time of a task T_i is known, we use the following formula to compute the arrival time of a subtask st_j^i of the task:

$$arrival\left(st_j^i\right) = arrival\left(T_i\right) + \sum_{k<j} dl\left(st_k^i\right) + dl\left(m_k^i\right)$$

We now illustrate the response time analysis of subtasks with a simple example. Assume the task set shown in Table 1 where the (relative) deadlines of subtasks and messages are computed from the end-to-end task deadline as described previously. Note that the sum of the subtask and message deadlines equals the end-to-end task deadline for all the tasks.

Table 1. An Example Task Set

Task	T_1	T_2	T_3	T_4
Type	aperiodic	aperiodic	periodic	periodic
Value	40	35	24	15
Period	--	--	20	35
Triggering Task	T_3	T_4	--	--
Subtasks and messages	$st_1^1, m_1^1, st_2^1, m_2^1$	$st_1^2, m_1^2, st_2^2, m_2^2, st_3^2, m_3^2$	$st_1^3, m_1^3, st_2^3, m_2^3$	st_1^4, m_1^4
End-to-end deadline	20	35	20	35
Subtask and message deadlines	6,2,10,2	10,2,11,2,8,2	6,1,12,1	29,6

[2] These assumptions will be relaxed in the next quarter of our work.

We consider the problem of determining the response time of subtask st_2^3 during the third period of task T_3. The arrival time of T_3 is assumed to be the beginning of the third period i.e., at time $t = (3-1)\times 20 = 40$. Then the arrival time of st_2^3 is computed by adding the relative deadlines of the subtask st_1^3 and the message m_1^3 to the arrival time of T_3. Thus, subtask st_2^3 is assumed to arrive at time $t = 40+6+1 = 47$.

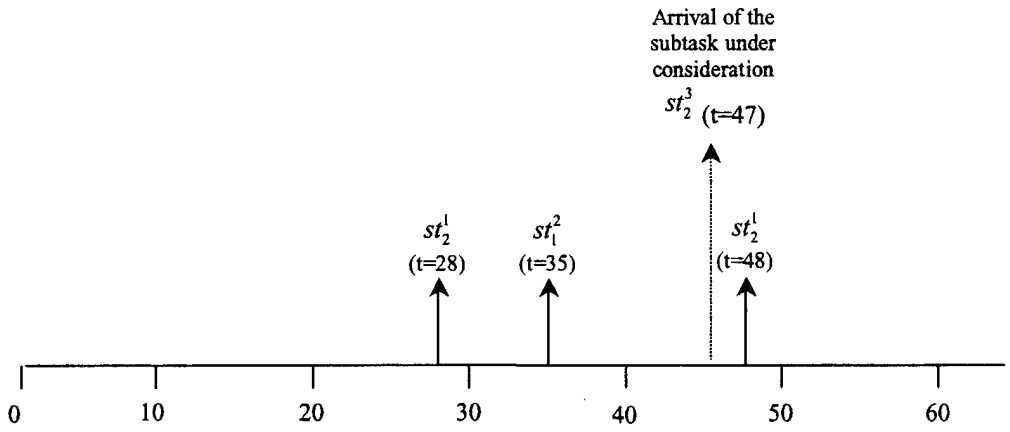


Figure 3. Arrival Times of Tasks for the Example Task Set

We will now consider a time window from $t = 0$ to the absolute end-to-end deadline of the parent task T_3 of the subtask under consideration (i.e., st_2^3). Since T_3 is assumed to arrive at $t = 40$, the absolute deadline is the summation of that arrival time plus the end-to-end deadline of T_3, which makes the time window 40 + 20 = 60. Throughout that time window, we will consider the arrival of all the subtasks of the higher value tasks on that processor if those subtasks have replicas that execute on that processor. We will consider the arrival of such subtasks throughout the time window described above from $t = 0$ to $t = 60$.

Suppose that the subtasks st_2^1 and st_1^2 of a higher value task have replicas that are running on the same processor as that of the subtask st_2^3. Further, assume that the replica of st_2^1 will be executing on this processor for the entire duration of the time window we are considering. Furthermore, assume that the replica of st_1^2 will be executing when the parent task T_2 is triggered by task T_4 during the first period. After the first period, the replica will be shut down.[3] We now compute the arrival times of the subtasks as described previously. Now throughout the time window, the arrival of the tasks will be as shown in Figure 3.

[3] This scenario is possible when we start and shut down replicas according to the expected load.

Once arrival times of subtasks are determined, the algorithm then predicts the subtask execution times using the adaptation functions to determine the expected load at each period. For the aperiodic tasks, this is the period of their triggering periodic tasks. The predicted execution times for the example task set are shown in Table 2 along with the response time analysis.

Table 2. Determining Subtask Response Times under LBESA

Iteration#	time	Arrival List	Process List	Sched. (LBESA) decision	Expected finishing time of the scheduled subtask	Next Event
0	0	$[st_2^1, 28, 8]$ $[st_2^1, 48, 9]$ $[st_1^2, 35, 7]$ $[st_2^3, 47, 6]$	Empty	N/A	N/A	Arrival $[st_2^1, 28, 8]$
1	28	$[st_2^1, 48, 9]$ $[st_1^2, 35, 7]$ $[st_2^3, 47, 6]$	$[st_2^1, 28, 8]$	st_2^1	Current time +remaining time= 28+8=36	Arrival $[st_1^2, 35, 7]$
2	35	$[st_2^1, 48, 9]$ $[st_2^3, 47, 6]$	$[st_2^1, 28, 1]$ $[st_1^2, 35, 7]$	st_2^1	35+1=36	Termination of st_2^1 at t=36
3	36	$[st_2^1, 48, 9]$ $[st_2^3, 47, 6]$	$[st_1^2, 35, 7]$	st_1^2	36+7=43	Termination of st_1^2 at t=43
4	43	$[st_2^1, 48, 9]$ $[st_2^3, 47, 6]$	Empty	N/A	N/A	Arrival $[st_2^3, 47, 6]$
5	47	$[st_2^1, 48, 9]$	$[st_2^3, 47, 6]$	st_2^3	47+6=53	Arrival $[st_2^3, 48, 6]$
6	48	Empty	$[st_2^1, 48, 9]$ $[st_2^3, 47, 5]$	st_2^1 (Preempt st_2^3)	48+9=57	Termination of st_2^1 at t=57
7	57	Empty	$[st_2^3, 47, 5]$	st_2^3	57+5=62	Termination of st_2^3 at t=62
8	62	Empty	Empty	N/A	N/A	N/A

Table 2 illustrates the procedure followed by the algorithm to determine the response time assuming that LBESA [Loc86] is used as the processor-scheduling algorithm. The algorithm uses two lists for determining the response times: *Arrival List* and *Process List*. The lists have three fields for each entry: (1) the subtask that is expected to arrive, (2) the time at which the arrival occurs, and (3) the remaining execution time of that subtask. The *Process List* represents the ready queue of the processor.

The algorithm first initializes the *Arrival List* to contain all the expected arrival events and it initializes the *Process List*. As illustrated in the table, the algorithm keeps a record of all the arrival events during the time window described above. The algorithm starts by processing the earliest event. For the example, it is the arrival of st_2^1 at $t = 28$. This entry is removed from the *Arrival List* and is placed in the *Process List*.

The algorithm now runs the local scheduler to determine which subtask from the ready queue will be selected to execute at that time. For the example, there is only one subtask in the ready queue and therefore will be selected by the scheduler for execution. The algorithm computes the expected finish time of the currently executing subtask by adding the current time to the remaining time of that subtask. The expected finish time is then compared to the earliest arrival time in the *Arrival List* to determine the earliest event, which is processed by the algorithm. When the finish time becomes earlier than the next arrival event, the algorithm terminates the current subtask and updates a local time variable with the finish time. Then, another scheduling decision is made to select the subtask from the ready queue. On the other hand, if the next arrival event is earlier than the finish time of the currently running subtask, the algorithm updates the local time variable with the arrival time and reduces the remaining time of the currently executing subtask by the amount of time that was elapsed since it was selected by the scheduler till the time of the next arrival. The next arrival event is removed from the *Arrival List* and is placed in the *Process List*. The local scheduler algorithm is then applied on the *Process List* to determine the new scheduling decision.

The process is repeated until the both lists become empty. The response time of each of the subtasks is computed as the difference between their arrival times and their finishing times. The response time of each subtask is then compared to its individual relative deadline. If all the subtasks satisfy their deadlines, the algorithm returns the response time of the subtask under consideration. If any subtask is found to misses its deadline, the algorithm returns a failure value. This means that replicating the subtask on that processor will affect the timeliness of the higher value tasks.

Table 3 shows the response times of the subtasks for the example task set.

Table 3. Response Times of Subtasks

subtask	Arrival time	Finishing time	Response time	Relative deadline
st_2^1 (first period of the triggering periodic task)	28	36	8	10
st_2^1 (second period of the triggering periodic task)	48	57	9	10
st_1^2	35	43	8	10
st_2^3	**47**	**62**	**15**	**12**

The last row of the table shows the expected performance of st_2^3 on the processor under consideration during the third period of its parent task T_3.

5.5 Estimating Message Communication Delays

In our proposed real-time network architecture, a message that is sent from a subtask to its successor is first transmitted from the host machine of the sender subtask to a router and then from the router to the destination subtask. Figure 4 illustrates the process.

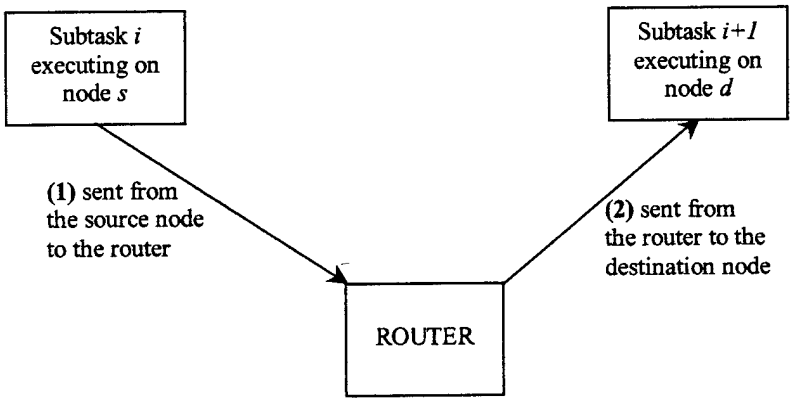


Figure 4. Routing of Messages on the Proposed Real-Time Network Architecture

The algorithm determines the communication delays of messages by assuming that no contentions occur. We will relax this assumption and investigate techniques for determining the communication delay under arbitrary contention conditions in the next quarter.

If no contention is assumed to occur, then the problem of estimating communication delay reduces to a very simple case in which the communication delay is nothing but transmission time. If l_s denote the link transmission speed, then the transmission delay for a message of size d is given by:

$$D_{trans}(d) = 2.\frac{d}{l_s}$$

Notice that we multiply by 2 because it takes the message two "hops" to reach the destination node from the source node. Also, notice that we assume that any processing done by the router is negligible compared to the transmission time since our initial assumption is that the message under consideration will experience no contention from the other messages.

5.6 Determining Number of Subtask Replicas and Their Processors

Recall the notation $T_i = [st_1^i, m_1^i, ..., st_p^i, m_p^i, st_q^i, m_q^i, st_r^i, m_r^i, ..., st_n^i, m_n^i]$ that denotes a task T_i consisting of n subtasks and n messages to be executed in series. Let subtask st_q^i be considered by the algorithm for replication. The algorithm first checks if the existing replica of the subtask (i.e., $st_{q,0}^i$) is enough for satisfying its individual deadline. This is done by determining the response time of the subtask (replica) on the processor p_i that was initially assigned to it. If the estimated response time is less than the individual deadline of st_q^i without affecting any of the subtasks of the higher value tasks, the algorithm assumes that the single replica of st_q^i is enough to satisfy its requirement and then considers the next subtask st_r^i.

If a single replica for st_q^i is not enough to satisfy its individual deadline or if executing st_q^i on that processor makes one or more of the subtasks of the higher value tasks miss their deadlines, the algorithm reduces the data size load of the subtask by replication. The algorithm considers a second replica for the subtask, which will reduce the data size load of the existing replica by half. To determine the processor for executing the second replica, the algorithm estimates the response time of the subtask for processing half the data size on each of the processors (excluding the processor p_i) as described in Section 5.4. The processor that gives the shortest response time, say processor p_j, is selected for executing the second replica. The algorithm now "backtracks" and computes the response time of the existing subtask replica $st_{q,0}^i$ on processor p_i under half the data size (since the second replica will now consume the other half of the data size). If the response time of both the replicas is found to be shorter than the subtask's individual deadline and the

execution of the replicas on the processors p_i and p_j is not found to affect the performance of the higher value tasks, then two replicas are considered to be sufficient by the algorithm.

If it is found that either of the two replicas will not satisfy the subtask deadline or will cause a subtask of a higher value task to miss its deadline, then the algorithm considers a third replica for the subtask and determines its response time. If the third replica is found to satisfy the subtask deadline without affecting the subtasks of the higher value tasks, the algorithm backtracks and ensures that the first two will not miss their deadlines. Thus, every time the algorithm considers adding a new replica, it backtracks and checks whether the old ones will be able to satisfy their deadlines without affecting the timeliness of higher value tasks.

BestEffortResourceManager(T,W)
/* T is the task set and W is the future time window */
1. Sort the task set in descending order according to task values;
2. For each task $i=1$ to $|T|$
/* from highest value task to lowest value task */
2.1 For each period $j=1$ to $\left\lceil \frac{W}{Period(T_i)} \right\rceil$
2.1.1 For each subtask $k=1$ to $|ST(T_i)|$
Replicate$\left(st_k^i, j, Adapt(T_i, j)\right)$;

Figure 5. The Best Effort Resource Management Algorithm

The process is repeated until each replica is able to satisfy the subtask deadline. Note that as the number of replicas increases, the workload share of each replica will be reduced. If the algorithm determines that executing the maximum number of replicas for a subtask (which is equal to the number of nodes in the system) does not satisfy the subtask deadline, it ignores the task by ignoring all its subtasks and considers the next task in the value-ordered list. The intuition for doing so is that by preserving the system resources from such infeasible tasks, we may be able to satisfy the deadlines of at least some subset of the lower value tasks and thereby increase the aggregate accrued value.

Figures 5, 6, and 7 shows the pseudo-code of the three main functions of the algorithm that determines the number of replicas for each subtask of a task set T for a future time window W. Note that this procedure for determining the number of replicas and their processors is the same no matter what local scheduling algorithm is used on the processors and at the router. The only difference lies in the function *LocalScheduler*()

that is called from the function *EstimateResponse*() for selecting the next subtask from the process ready queue. This function will vary depending upon the scheduling algorithm.

Replicate(s,i,l)

/* s is the subtask, i is the period, and l is the anticipated load */

1. if *EstimateResponse*$\left(s,i,prr(s_0),l\right) < dl(s)$ /* s_0 is the first replica of s */

    1.1 return SUCCESS; /* if the first replica is enough, return */

2. $PT = PR - pr(s)$; /* PT is the set of all processors that are not already executing s */

3. if $PT = \varnothing$

    3.1 return FAILURE;

4. For each processor $q \in PT$

    4.1 *ResponseTime* = *EstimateResponse*$\left(s,i,q,\frac{l}{|pr(s)|+1}\right)$;

    /* $|pr(s)+1|$ is the number of replicas after increasing it by one and $\frac{l}{|pr(s)+1|}$ is the new load */

    4.2 Determine the minimum response time and the corresponding processor. Save the value of the minimum response time in the local variable *MinResponse* and save the corresponding processor ID in p_{min};

5. $PT = PT - \{p_{min}\}$;

6. $pr(s) = pr(s) \cup \{p_{min}\}$;

7. if *MinResponse* $> dl(s)$

    6.1 return to step 3;

8. For each processor $q \in pr(st_j^i) - \{p_{min}\}$

    8.1 if *EstimateResponse*$\left(s,i,q,\frac{l}{|pr(s)|}\right) > dl(s)$

        8.1.1 return to step 3;

9. return SUCCESS;

Figure 6. The Best Effort Resource Management Algorithm – function *Replicate*()

EstimateResponse(s, p, q, l)

/* s is the subtask, p is the period number, q is the processor, l is the workload */

1. $ArrivalEvents[\ ]=\emptyset$; $Process[\ ]=\emptyset$;

/* Lists $ArrivalEvents[\ ]$ and $Process[\ ]$ have 3 fields: time, subtask, remaining_time */

2. $x = ParentTaskNumber(s)$; $y = SubtaskNumber(y, x)$; /* y is subtask # of s within its parent task x */

3. if $type(s) = PERIODIC$

   3.1 $ArrivalTime(s) = (p-1)\times period(T_x) + \sum_{r<y}\left(dl(st_r^x) + dl(m_r^x)\right)$;

   3.2 $StopTime = p \times period(T_x)$;

4. else

   4.1 $ArrivalTime(s) = p \times period(T_x) + \sum_{i<y}\left(dl(st_i^x) + dl(m_i^x)\right)$;

   4.2 $StopTime = (p+1)\times period(T_x)$;

5. $ArrivalEvents = ArrivalEvents \cup \left[st_y^x, ArrivalTime(s), eex(s,l)\right]$;

6. For each task i = 1 to $x-1$ /* consider all tasks that have higher values */

   6.1 For each period $j = 1$ to $\left\lceil \frac{StopTime}{period(T_i)} \right\rceil$

      6.1.1 For each subtask $k = 1$ to $|ST(T_i)|$

         6.1.1.1 if subtask st_k^i has a replica executing on processor q during period j

            1. $ArrivalTime(st_k^i) = p \times period(T_i) + \sum_{r<k}\left(dl(st_r^i) + dl(m_r^i)\right)$;

            2. if type$(st_k^i) = APERIODIC$

               $ArrivalTime(st_k^i)$ += $period(T_i)$;

            3. $ArrivalEvents = ArrivalEvents \cup \left[st_k^i, ArrivalTime(s), eex(s,l)\right]$;

7. $TerminationTime = \infty$; $SchedulerTime = 0$; $CurrentProcess = 1$;

8. Determine the earliest event $ArrivalEvents[k]$ from $ArrivalEvents$;

9. If $ArrivalEvents[k].time < TerminationTime$

   9.1 $Process[CurrentProcess].remaining\_time\ -= ArrivalEvents[k].time - t$;

   9.2 $t = ArrivalEvents[k].time$;

   9.3 Add $ArrivalEvents[k]$ to $Process[\ ]$ and remove $ArrivalEvents[k]$ from $ArrivalEvents[\ ]$;

   9.4 $CurrentProcess = LocalScheduler(Process[\ ])$;

   9.5 $TerminationTime = t + Process[CurrentProcess].remaining\_time$;

   9.6 return to step 8;

10. else

   10.1 $t = TerminationTime$;

   10.2 if $Process[CurrentProcess].subtask = s$

      10.2.1 $ResponseTime = t - ArrivalTime(s)$;

   10.3 if $t > Process[CurrentProcess].time + dl(Process[CurrentProcess].subtask)$

      /*one of the subtasks of the higher value tasks missed its deadline*/

      10.3.1 return ∞;

   10.4 Remove $Process[CurrentProcess]$ from $Process[\ ]$;

   10.5 return to step 8;

11. return $ResponseTime$;

Figure 7. The Best Effort Resource Management Algorithm – function *EstimateResponse*()

5.7 Experimental Evaluation

We evaluate the performance of the algorithm using a combination of benchmarking and simulation. We summarize the simulation parameters in Section 5.7.1. Section 5.7.2 presents the performance results of the algorithm.

5.7.1 Baseline Parameters of the Simulation

Table 4 summarizes the baseline parameters of the simulation study. The baseline parameters are derived from the real-time benchmark that has resulted from our past work [SWR99]. The structure of the periodic and aperiodic tasks used here corresponds directly to that of the benchmark.

Table 4. Baseline Parameters of the Simulation

Number of periodic Tasks	20
Number of aperiodic Tasks	20
Number of nodes	15
Range of number of subtasks per task	1-8
Size of a data item (track)	80 bytes
Range of periods lengths of periodic tasks	200 ms-4 sec
Relative end-to-end deadline for periodic and aperiodic tasks	same as the corresponding periods
Range of task absolute values	200-4500
Simulation time	10 sec

The coefficients of the regression equation presented in Section 5.4 (equation 3) for two application subtasks of the benchmark are shown in Table 5.

Table 5. Coefficients of the Execution Latency Regression Equation

Subtask number	a_1	a_2
3	0.090741	1.6519629
5	0.093583	0.4844743

5.7.2 Simulation Results

We evaluate the performance of the resource management algorithm under three scheduling algorithms that are used at all the processors and at the router. These include DASA [Cla90], LBESA [Loc86], and EDF [SSRB98]. The adaptation functions used in

the simulation study are increasing ramp functions. In other words, the task loads are anticipated to monotonically increase with time. We consider the performance metrics of (1) overall value accrued which is measured as the sum of the value accrued by the execution of all the tasks and (2) the missed deadline ratio.

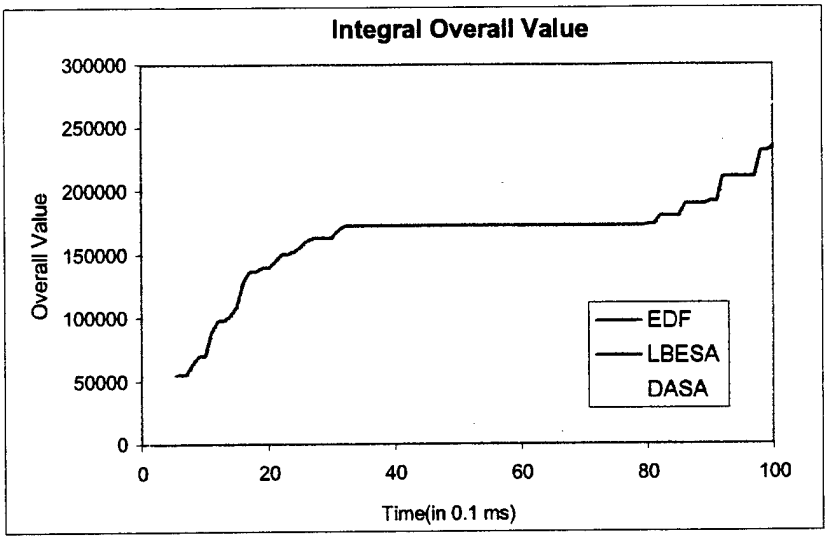


Figure 8. Overall Value Accrued by the Best-Effort Resource Management Algorithm under DASA, LBESA, and EDF

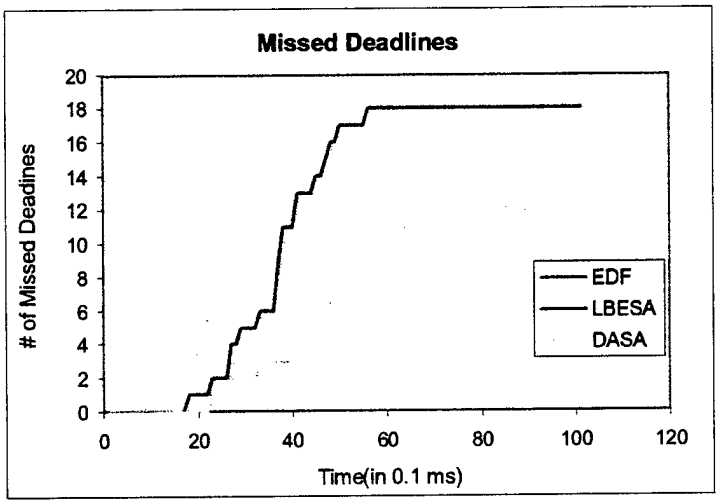


Figure 9. Missed Deadline Ratio of the Best-Effort Resource Management Algorithm under DASA, LBESA, and EDF

Figure 8 shows the overall value accrued by the best-effort resource management algorithm under each of the three scheduling algorithms during the experiments for every 100 milliseconds. Figure 9 shows the corresponding missed deadline ratio.

From the figures, we observe that our best effort resource management algorithm under DASA and LBESA produce higher overall accrued value and lower missed deadline ratios than that under EDF. The performance of the resource management algorithm under DASA and LBESA are also found to be exactly the same (the plots overlap for all data points for the two algorithms).

Thus, the experimental results illustrate the superiority of our best effort resource management algorithm under the DASA and LBESA best-effort schedulers. Note that in [Cla90], DASA is shown to be better than LBESA and EDF for both the performance metrics of overall value accrued and missed deadline ratios. However, this result is shown in [Cla90] for a single processor system under deterministic conditions, where subtask execution times are known with absolute certainty and do not fluctuate at run-time.

6. Publication, Presentation, and Related Activities

Our publication, presentation, and other related activities during this quarter includes the following:

Publications

1. B. Ravindran, "Heuristic Algorithms for Adaptive Communication in Soft Real-Time Distributed Systems," *IEEE/ACM Transactions on Networking*, Submitted January 2001, (under review)
2. B. Ravindran and T. Hegazy, "A Predictive Algorithm for Adaptive Resource Management of Periodic Tasks in Asynchronous Real-Time Distributed Systems," *IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, April 2001, Accepted for publication, To appear.
3. B. Ravindran, T. Hegazy, and P. Kachroo, "Adaptive Resource Management in Asynchronous Real-Time Distributed Systems Using Feedback Control Functions," *The Fifth IEEE International Symposium on Autonomous Decentralized Systems*, March 2001, Accepted for publication, To appear.
4. B. Ravindran, R. Devarasetty, and B. Shirazi, "Heuristic Algorithms for Adaptive Resource Management of Periodic Tasks in Soft Real-Time Distributed Systems," *Journal of Parallel and Distributed Computing*, Initial submission September 98, Revised submission under preparation.
5. B. Ravindran, P. Kachroo, and T. Hegazy, "Intelligent Feedback Control–based Adaptive Resource Management for Asynchronous, Decentralized Real-Time Systems," *IEEE Transactions on Systems, Man, And Cybernetics*, Initial submission July 2000, Revised submission under preparation.

6. B. Ravindran and T. Hegazy, "Proactive Resource Management of Aperiodic Computations in Asynchronous, Real-Time Distributed Systems," *IEEE Transactions on Computers*, (under preparation for submission)

7. Planned Efforts for the Next Quarter

Our planned activities for the next quarter includes the following:

1. Design *efficient* best-effort resource management algorithms for proactive resource management by:
   (a) developing fast and efficient techniques for determining subtask response times under best effort scheduling algorithms for arbitrary time intervals
   (b) developing fast and efficient variants of DASA and other best effort scheduling algorithms
2. Design *decentralized* best-effort resource management algorithms for proactive resource management
3. Implement adaptive resource management techniques in a real-time middleware and experimentally evaluate the performance of the techniques using a real-time benchmark
4. Design *adaptive communication techniques for IEEE 802.3* (Ethernet) by
   (a) developing novel real-time Ethernet architectures
   (b) developing best-effort routing algorithms for adapting end-to-end application messages on such architectures for accommodating workload uncertainties. The best effort routing algorithms will seek to provide the best aggregate timeliness to the application—as specified by the application—that is possible under the current application and resource situation.

References

[ABR+93] N. Audsley, A. Burns, M. Richardson, K. Tindell, A. Wellings, "Applying New Scheduling Theory to Static Priority Pre-emptive Scheduling," *Software Engineering Journal*, Vol. 8, No. 5, pages 284-292, September 1993.

[Cla90] R. K. Clark, "Scheduling Dependent Real-Time Activities," CMU-CS-90-155 (*Ph.D. Thesis*), Department of Computer Science, Carnegie Mellon University, 1990.

[IEEE802a] IEEE's 802.1p Standard, Available at: http://standards.ieee.org/catalog/IEEE802.1.html

[IEEE802b] IEEE's 802.1p Standard (discussion), Available at: http://www.nwfusion.com/news/tech/0907tech.html

[Jen92] E. D. Jensen, "Asynchronous Decentralized Real-Time Computer Systems," Book Chapter, *Real-Time Computing, Proceedings of the NATO Advanced Study Institute*, St. Martin, Springer-Verlag, October, 1992.

[Jen93] E. D. Jensen, "A Timeliness Model for Asychronous Decentralized Computer Systems," *Proceedings of the IEEE International Symposium on Autonomous Decentralized Systems*, 1993

[Jen99] E. D. Jensen, "Adaptive Real-Time Distributed Computer Systems," Available at: http://www.realtime-os.com/

[KG97] B. Kao and H. Garcia-Molina, "Deadline Assignment in a Distributed Soft Real-time System," *IEEE Transactions on Parallel and Distributed Systems*, Volume: 8 Issue: 12, pages 1268 –1274, December 1997.

[Loc86] C. D. Locke, "Best-Effort Decision Making for Real-Time Scheduling," CMU-CS-86-134 (*Ph.D. Thesis*), Department of Computer Science, Carnegie Mellon University, 1986.

[Mills95] D. L. Mills, "Improved Algorithms for Synchronizing Computer Network Clocks," *IEEE/ACM Transactions on Networks*, pages 245 – 254, June 1995.

[MS00] A. S. Matveev and A. V. Savkin, "Qualitative Theory of Hybrid Dynamical Systems," Birkhauser, 2000.

[MV+96] R. C. Metzger, B. VanVoorst, L. S. Pires, R. Jha, W. Au, M. Amin, D. A. Castanon, and V. Kumar, "C3I Parallel Benchmark Suite – Introduction and Preliminary Results," *Supercomputing*, 1996.

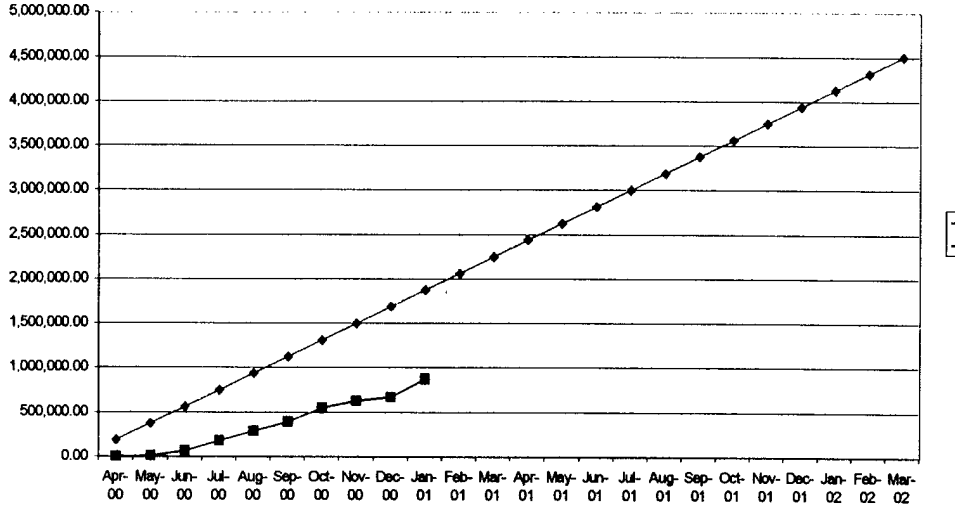
[Quo97] DARPA ITO, "Quorum," Available at http://www.ito.darpa.mil/research/quorum/projlist.html, August 1997.

[SWR99] B. Shirazi, L. R. Welch, and B. Ravindran, "DynBench: A Benchmark Suite for Dynamic Real-Time Systems," *Journal of Parallel and Distributed Computing Practices*, 1999, Accepted for publication, To appear. (Available at http://www.ee.vt.edu/~binoy/papers.html)

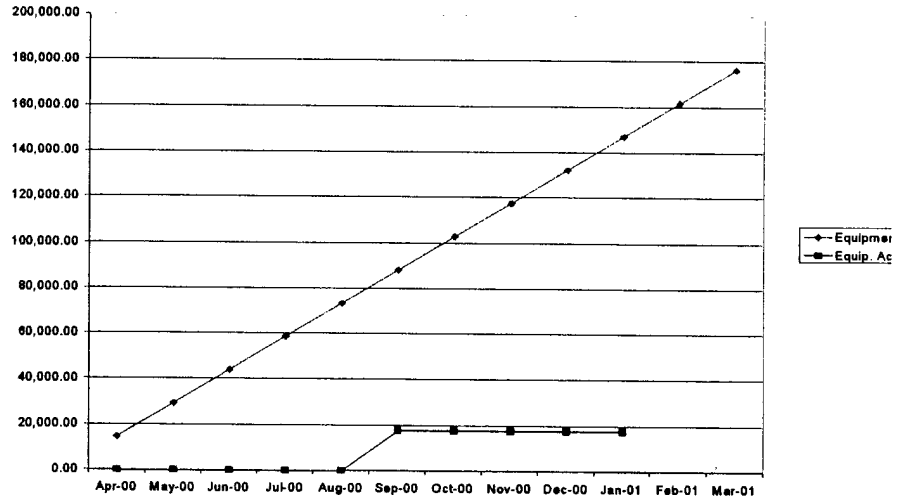
[WRSB98] L. R. Welch, B. Ravindran, B. A. Shirazi, and C. Bruggeman, "Specification and Modeling of Dynamic, Distributed Real-Time Systems," *Proceedings of The 19th IEEE Real-Time Systems Symposium*, pages 72 - 81, December 1998.

Financial Status

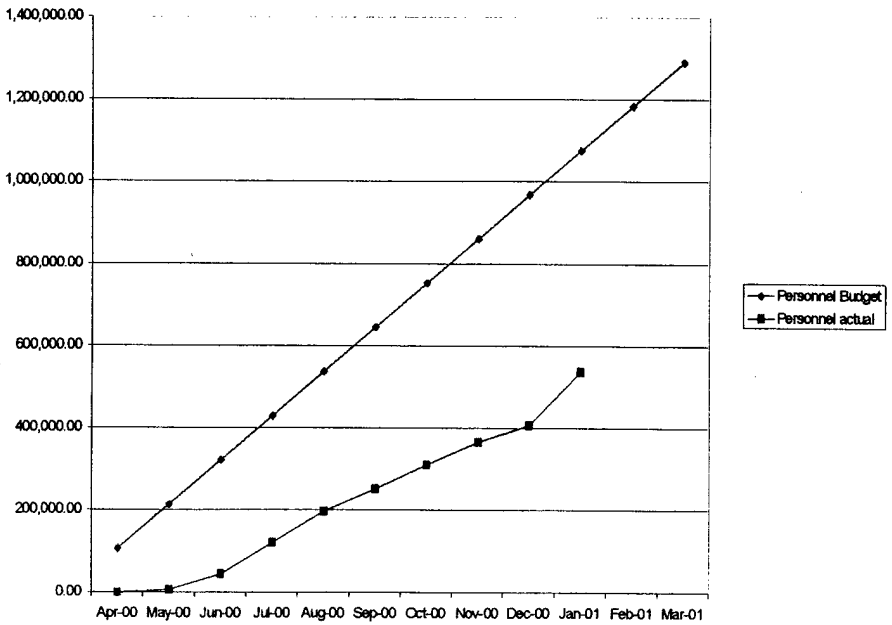
Total Project Budget/Actual



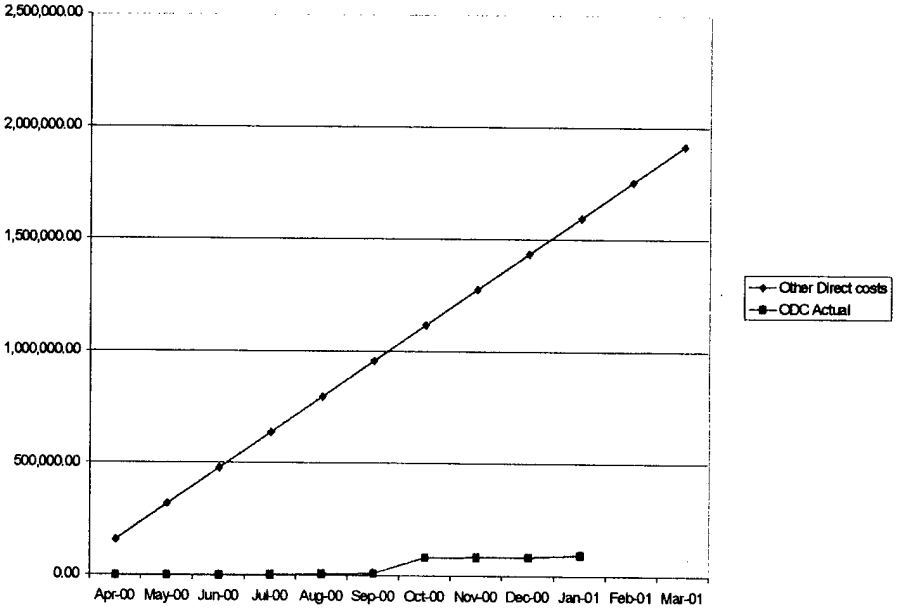
Equipment Budget vs. Actual



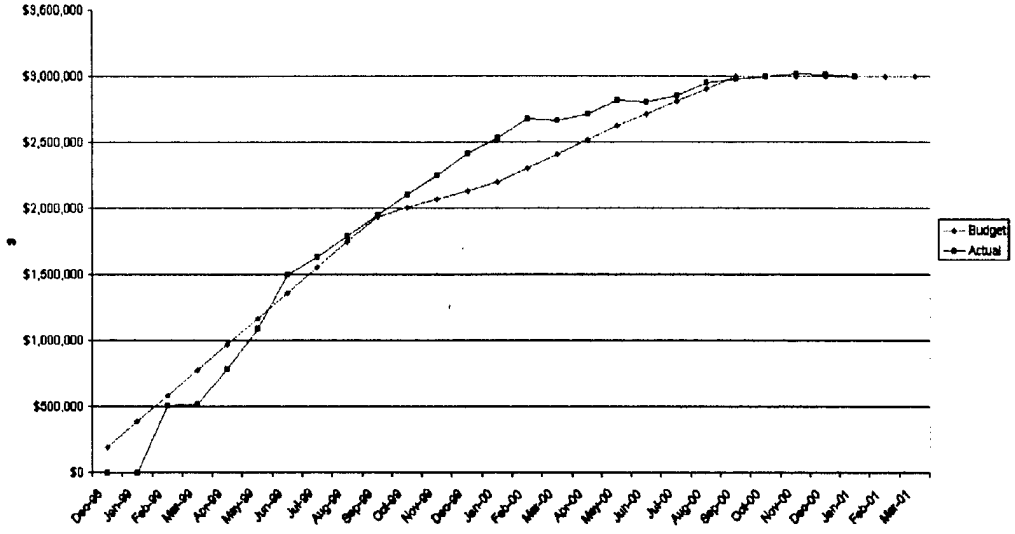
Personnel Budget vs. Actual



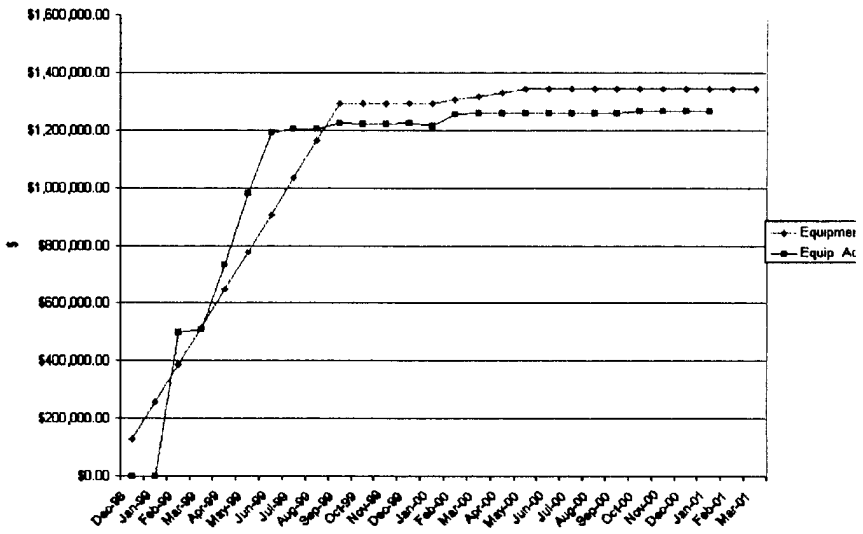
Other Direct Costs vs. Budget



NAVCITI Total Project Budget vs. Actual Expenditures + Commitments

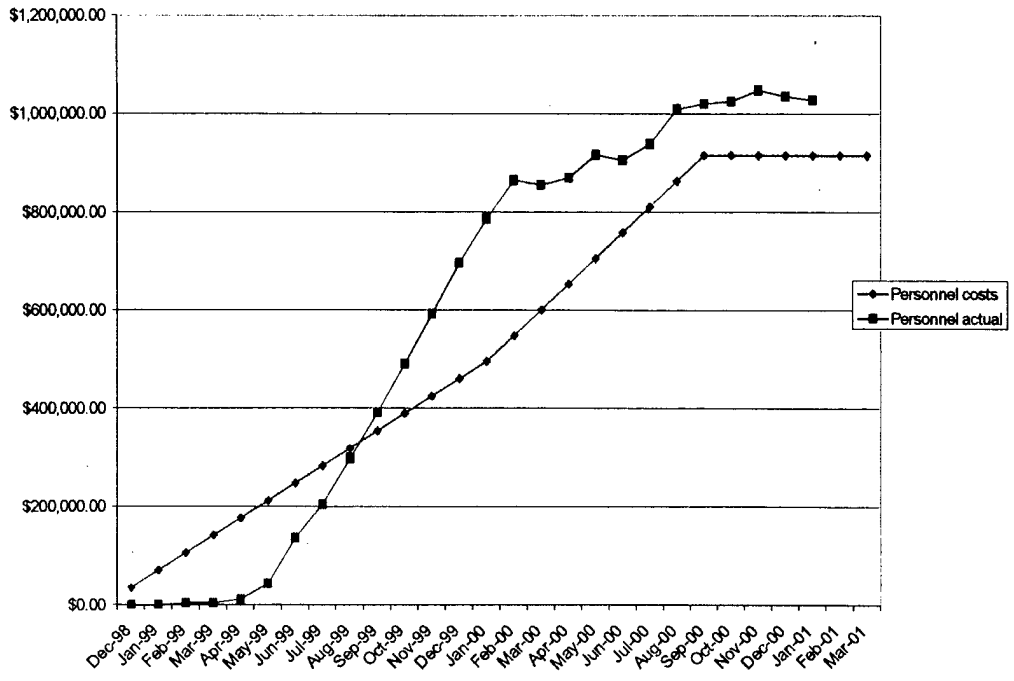


NAVCIITI Equipment & Software Purchases vs. Actual



100

NAVCIITI Personnel Budget vs. Actual



NAVCITI Other Direct Costs vs. Actual